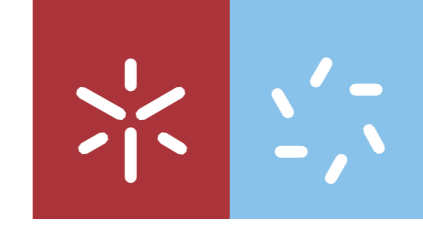


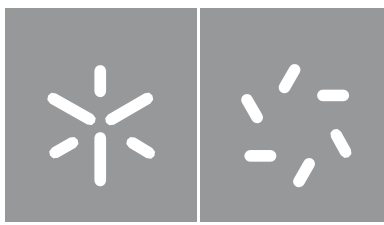


**Acidentes de Trabalho – Tábua de
Mortalidade Geracional**

Margarida Castro de Oliveira

Universidade do Minho
Escola de Ciências





Universidade do Minho
Escola de Ciências

Margarida Castro de Oliveira

Acidentes de Trabalho – Tábua de Mortalidade Geracional

Dissertação de Mestrado
Mestrado em Estatística para
Ciência de Dados

Trabalho efetuado sob a orientação da

**Professora Arminda Manuela Andrade Pereira
Gonçalves**

e do

Dr. Filipe Alexandre Carvalho Teixeira Gonçalves

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



**Atribuição
CC BY**

<https://creativecommons.org/licenses/by/4.0/>

Agradecimentos

A realização desta dissertação de fim de mestrado contou com a ajuda de diversas pessoas em que, sem as mesmas, não se teria tornado realidade e às quais estarei eternamente grata.

Primeiramente, à minha orientadora Professora Arminda Manuela Gonçalves, pela orientação, disponibilidade e pelo conhecimento que me transmitiu. Sem as suas opiniões e críticas a realização deste relatório não seria possível.

Seguidamente, ao meu orientador da Ageas, Dr. Filipe Gonçalves, e ao Dr. Luís Maranhão, evidenciando o seu total desempenho no presente tema.

A proatividade, exigência e rigor de todos fez com que me ajudassem a solucionar inúmeros desafios que foram surgindo. Por isso agradeço, pois, mesmo sendo um estágio remoto, nunca falharam no auxílio e acompanhamento das minhas tarefas.

Por fim, e não menos importante, quero agradecer à minha família, aos meus pais pelo esforço enorme que fizeram para que eu pudesse tirar um curso universitário, à minha irmã pela força que sempre me transmitiu e aos meus avós pelos pensamentos positivos que me demonstraram ao longo da vida.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

“Sê todo em cada coisa.

Põe quanto és no mínimo que fazes.”

Odes de Ricardo Reis, Fernando Pessoa.

Resumo

A esperança de vida tem vindo a aumentar ao longo dos anos, sendo por isso necessário atualizar as tábuas de mortalidade da população portuguesa. Nestas constam indicadores bastante importantes como a taxa de mortalidade e a esperança média de vida e, através de algumas funções que constam nas mesmas, é também possível calcular as taxas vitalícias dos pensionistas de modo a realizar um cálculo mais correto das responsabilidades da companhia.

O objetivo deste trabalho é construir tábuas geracionais, que modelem o comportamento da população portuguesa, e pretende-se também desenvolver modelos que permitam estimar os fatores a serem utilizados para o cálculo das pensões vitalícias no ramo de Acidentes de Trabalho. Inicialmente os dados a modelar são referentes à população portuguesa e, em seguida, é efetuado o ajustamento aos dados específicos do mercado segurador.

Com o objetivo de construir as tábuas de mortalidade, inicialmente é necessário modelar a mortalidade, utilizando métodos como o Método Lee-Carter e o Modelo de Poisson (ou Extensão do Método Lee-Carter). Para modelar e prever as séries temporais subjacentes às taxas de mortalidade, são utilizados modelos de previsão de séries temporais paramétricos (Modelação ARIMA) e não paramétricos (Método de Holt).

Por fim, são modeladas as taxas de mortalidade da população da companhia, de modo a comparar as taxas obtidas na população portuguesa e na população da companhia, para entender se o comportamento dos pensionistas é semelhante ao da população portuguesa. Esta comparação é feita através do cálculo de medidas de avaliação, nomeadamente o EQM, o REQM, o EPAM, o EAM e a estatística U de Theil.

Palavras-Chave: Acidentes de Trabalho, Tábuas de Mortalidade, INE, ASF, Método Lee-Carter, ARIMA, Modelo de Poisson, Método de Holt.

Abstract

Life expectancy has been increasing over the years and it is therefore necessary to update the mortality tables for the Portuguese population. These tables contain very important indicators such as the mortality rate and the life expectancy and through some functions that are included in them, it is also possible to calculate the lifetime rates of pensioners in order to make a more accurate calculation of the company's liabilities.

The objective of this work is to build generational tables that model the behavior of the Portuguese population and also to develop models that allow estimating the factors to be used to calculate the life annuities in the Work Accidents branch. Initially, the data to be modeled refers to the Portuguese population and, then, the adjustment is made to the specific data of the insurance market.

In order to construct the life tables, it is initially necessary to model mortality, using methods such as the Lee-Carter Method and the Poisson Model (or Extension of the Lee-Carter Method). To model and forecast the time series underlying the mortality rates, both parametric (ARIMA Modeling) and non-parametric (Holt's Method) time series forecasting models are used.

Finally, the mortality rates of the company's population are modeled, in order to compare the rates obtained in the Portuguese population and in the company population, to understand if the behavior of the pensioners is similar to that of the Portuguese population. This comparison is made by calculating evaluation measures, namely the MSE, the RMSE, the MAPE, the MAE and Theil's U statistic.

Keywords: Work Accidents, Mortality Tables, INE, ASF, Lee-Carter Method, ARIMA, Poisson Model, Holt Method.

Índice

1	Introdução	1
1.1	Companhia de Seguros Ageas	2
1.1.1	Ramo Acidentes de Trabalho	3
1.2	Objetivos do Trabalho	3
1.3	Estrutura do Trabalho	4
2	Tábuas de Mortalidade	5
2.1	Classificação das Tábuas de Mortalidade	6
2.2	Projeção de Tábuas de Mortalidade	8
3	Método Lee-Carter e Modelo de Poisson (Extensão do Método Lee-Carter)	11
3.1	Método Lee-Carter	11
3.2	Extensão do Método Lee-Carter (Modelo de Poisson)	13
3.3	Ajustamento do Modelo Demográfico	13
3.3.1	Método Lee-Carter	13
3.3.2	Extensão do Método Lee-Carter (ou Modelo de Poisson)	16
3.4	Modelação do Nível Geral de Mortalidade	19
3.5	Limitações do método Lee-Carter e novos desenvolvimentos	19
4	Séries Temporais	22
4.1	Conceito de Série Temporal	22
4.2	Processos Estocásticos	24
4.2.1	Processos Estocásticos Estacionários	25
4.2.2	Funções de Autocovariância, Autocorrelação e Autocorrelação Parcial	26

4.2.3	Ruído Branco	28
4.2.4	Estimação dos parâmetros caracterizadores dos processos estacionários	29
4.3	Processos Estocásticos Não Estacionários	29
4.3.1	Transformações para a estacionariedade	30
4.3.2	Análise de estacionariedade	34
5	Modelos de Previsão em Séries Temporais	38
5.1	Metodologia Box-Jenkins	38
5.2	Modelos de Processos Estacionários	39
5.2.1	Processo Autorregressivo de ordem p , $AR(p)$	40
5.2.2	Processo de Médias Móveis (MA)	41
5.2.3	Processo Autorregressivo e de Médias Móveis, $ARMA(p, q)$	42
5.3	Modelos de Processos Não Estacionários	44
5.3.1	Processo Autorregressivo Integrado de Médias Móveis (ARIMA)	44
5.3.2	Processo Autorregressivo Integrado de Médias Móveis Sazonal (SARIMA)	45
5.4	Etapas da Metodologia Box-Jenkins	47
5.5	Análise dos Resíduos	49
5.5.1	Distribuição Normal	49
5.5.2	Não Correlação	50
5.5.3	Média Nula e Variância Constante	50
5.6	Seleção dos Modelos	51
5.6.1	Crítério de Informação Akaike	51
5.6.2	Crítério de Informação Bayesiano	52
5.7	Previsão Pontual e Intervalar	52
5.8	Métodos de Alisamento Exponencial	53
5.8.1	Método de Alisamento Exponencial Simples	55
5.8.2	Método Linear de Holt	56
5.8.3	Método de Holt-Winters	58
5.8.4	Outros Métodos de Alisamento Exponencial	60
5.8.5	Intervalos de Previsão	62
5.9	Medidas de Avaliação	63

6	Método de Lee-Carter e Modelo de Poisson para a População Portuguesa	65
6.1	Análise Descritiva dos Dados	65
6.1.1	Caso 1: Sexo Masculino	66
6.1.2	Caso 2: Sexo Feminino	69
6.1.3	Caso 3: Sem Distinção de Sexo	72
6.2	Aplicação do Método Lee-Carter	76
6.3	Modelação do Nível Geral de Mortalidade - Modelação ARIMA	78
6.3.1	Identificação do Modelo	79
6.3.2	Avaliação do Modelo	86
6.4	Modelação do Nível Geral de Mortalidade - Método de Holt	94
6.5	Extensão do Método Lee-Carter ou Modelo de Poisson	97
6.6	Modelação do Nível Geral de Mortalidade - Modelação ARIMA	100
6.6.1	Identificação do Modelo	100
6.6.2	Avaliação do Modelo	108
6.7	Modelação do Nível Geral de Mortalidade - Método de Holt	112
6.8	Método Lee-Carter vs. Extensão do Método Lee-Carter	115
6.8.1	Modelação ARIMA	115
6.8.2	Método de Holt	126
6.8.3	Modelação ARIMA vs. Método de Holt	128
7	Tábuas de Mortalidade da População Portuguesa	129
7.1	Cálculo das Pensões Vitalícias	133
8	Aplicação aos Dados da Companhia	135
8.1	Análise Descritiva dos Dados	135
8.1.1	Caso 1: Sexo Masculino	135
8.1.2	Caso 2: Sexo Feminino	138
8.1.3	Caso 3: Sem Distinção de Sexo	140
8.2	Tratamento de Dados - Médias Móveis e Médias Móveis Ponderadas	144
9	Conclusão	151
9.1	Sugestões para Trabalho Futuro	152

Apêndices	157
A Análise Descritiva dos Dados Portugueses	158
B Estimação dos Parâmetros (Método Lee-Carter)	173
C Estimação dos Parâmetros (Extensão do Método Lee-Carter ou Modelo de Poisson)	194
D Tábuas de Mortalidade	215
E Cálculo das Taxas Vitalícias	238
F Análise Descritiva dos Dados da Companhia	261
G Comparação das taxas de mortalidade da população portuguesa e da população da companhia	264

Índice de Figuras

2.1	Excerto da tábua de mortalidade Portuguesa de 2018-2020, para ambos os sexos, realizada pelo Instituto Nacional de Estatística (INE).	10
4.1	Simulação de um ruído branco e respetivas FAC e FACP empíricas.	28
4.2	Simulação de um modelo de passeio aleatório e as suas diferenças de 1. ^a ordem.	32
4.3	Simulação de um modelo de passeio aleatório com <i>drift</i> e as suas diferenças de 1. ^a ordem.	33
5.1	Simulação de um processo autorregressivo e as respetivas FAC e FACP empíricas.	41
5.2	Simulação de um processo de médias móveis e as respetivas FAC e FACP empíricas.	42
5.3	Simulação de um processo autorregressivo e de médias móveis e as respetivas FAC e FACP empíricas.	43
5.4	Simulação de um processo autorregressivo e de médias móveis integrado, ARIMA(2, 1, 1) e respetivas FAC e FACP empíricas.	45
5.5	Simulação de um processo autorregressivo e de médias móveis integrado sazonal, SARIMA(2, 1, 1)(1, 1, 1) ₁₂ e respetivas FAC e FACP empíricas.	46
6.6	Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sexo Masculino.	66
6.7	<i>Boxplots</i> do número de mortes nos anos 1940-2020 - Sexo Masculino.	67
6.8	<i>Zoom</i> dos <i>boxplots</i> do número de mortes nos anos 1940-2020 - Sexo Masculino.	67
6.9	Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sexo Masculino.	68
6.10	<i>Boxplots</i> do número de expostos ao risco nos anos 1940-2020 - Sexo Masculino.	69
6.11	Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sexo Feminino.	70
6.12	<i>Boxplots</i> do número de mortes nos anos 1940-2020 - Sexo Feminino.	70
6.13	<i>Zoom</i> dos <i>boxplots</i> do número de mortes nos anos 1940-2020 - Sexo Feminino.	71
6.14	Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sexo Feminino.	71
6.15	<i>Boxplots</i> do número de expostos ao risco nos anos 1940-2020 - Sexo Feminino.	72

6.16	Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sem Distinção de Sexo.	73
6.17	<i>Boxplots</i> do número de mortes nos anos 1940-2020 - Sem Distinção de Sexo.	73
6.18	<i>Zoom</i> dos <i>boxplots</i> do número de mortes nos anos 1940-2020 - Sem Distinção de Sexo.	74
6.19	Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sem Distinção de Sexo.	75
6.20	<i>Boxplots</i> do número de expostos ao risco nos anos 1940-2020 - Sem Distinção de Sexo.	75
6.21	Estimativas do nível médio das taxas de mortalidade, α_x , ao longo das idades 0-100 anos.	77
6.22	Estimativas das oscilações nas taxas de mortalidade, β_x , ao longo das idades 0-100 anos.	77
6.23	Estimativas do nível geral de mortalidade, k_t , nos anos 1940-2020.	78
6.24	Série temporal do parâmetro k_t no período de treino (1940-2010) - Sexo Masculino.	79
6.25	Série temporal do parâmetro k_t no período de treino (1940-2010) - Sexo Feminino.	80
6.26	Série temporal do parâmetro k_t no período de treino (1940-2010) - Sem Distinção de Sexo.	80
6.27	Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Masculino.	80
6.28	Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Feminino.	81
6.29	Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sem Distinção de Sexo.	81
6.30	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sexo Masculino.	82
6.31	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sexo Feminino.	82
6.32	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.	83
6.33	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sexo Masculino.	83
6.34	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sexo Feminino.	84
6.35	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.	84

6.36	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.	87
6.37	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.	88
6.38	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.	88
6.39	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sexo Feminino.	89
6.40	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sem Distinção de Sexo.	90
6.41	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.	91
6.42	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.	92
6.43	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.	92
6.44	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.	93
6.45	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.	93
6.46	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.	95
6.47	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.	95

6.48	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.	96
6.49	Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sexo Masculino.	98
6.50	Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sexo Feminino.	98
6.51	Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sem Distinção de Sexo.	98
6.52	Estimativas do parâmetro k_t nos anos 1940-2020 - Sexo Masculino.	99
6.53	Estimativas do parâmetro k_t nos anos 1940-2020 - Sexo Feminino.	99
6.54	Estimativas do parâmetro k_t nos anos 1940-2020 - Sem Distinção de Sexo.	99
6.55	Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sexo Masculino.	100
6.56	Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sexo Feminino.	101
6.57	Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sem Distinção de Sexo.	101
6.58	Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Masculino.	102
6.59	Função de autocorrelação e de autocorrelação parcial a série temporal do nível geral de mortalidade - Sexo Feminino.	102
6.60	Função de autocorrelação e de autocorrelação parcial a série temporal do nível geral de mortalidade - Sem Distinção de Sexo.	102
6.61	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sexo Masculino.	103
6.62	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sexo Feminino.	104
6.63	Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.	104
6.64	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sexo Masculino.	105
6.65	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sexo Feminino.	105

6.66	Série do nível geral de mortalidade k_t após a 2ª diferenciação (d=2) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.	106
6.67	Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.	109
6.68	Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.	110
6.69	Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.	110
6.70	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.	111
6.71	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.	111
6.72	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.	112
6.73	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.	113
6.74	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.	114
6.75	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.	114
6.76	Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.	117
6.77	Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.	118

6.78	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sexo Masculino.	119
6.79	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.	119
6.80	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.	122
6.81	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.	122
6.82	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.	124
6.83	Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.	125
6.84	Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.	126
7.85	Evolução da taxa de mortalidade - Sexo Masculino.	130
7.86	Evolução da taxa de mortalidade - Sexo Feminino.	130
7.87	Evolução da taxa de mortalidade - Sem Distinção de Sexo.	131
7.88	Evolução da esperança de vida - Sexo Masculino.	132
7.89	Evolução da esperança de vida - Sexo Feminino.	132
7.90	Evolução da esperança de vida - Sem Distinção de Sexo.	133
8.91	Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Masculino.	136
8.92	<i>Boxplots</i> do Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Masculino.	136
8.93	Número de Mortes nos Anos 2007-2021 - Sexo Masculino.	137
8.94	<i>Boxplots</i> do Número de Mortes nos Anos 2007-2021 - Sexo Masculino.	137
8.95	Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Feminino.	138
8.96	<i>Boxplots</i> do Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Feminino.	139
8.97	Número de Mortes nos Anos 2007-2021 - Sexo Feminino.	139
8.98	<i>Boxplots</i> do Número de Mortes nos Anos 2007-2021 - Sexo Feminino.	140

8.99	Número de Expostos ao Risco nos Anos 2007-2021 - Sem Distinção de Sexo.	141
8.100	<i>Boxplots</i> do Número de Expostos ao Risco nos Anos 2007-2021 - Sem Distinção de Sexo.	141
8.101	Número de Mortes nos Anos 2007-2021 - Sem Distinção de Sexo.	142
8.102	<i>Boxplots</i> do Número de Mortes nos Anos 2007-2021 - Sem Distinção de Sexo.	142
8.103	Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sexo Masculino.	145
8.104	Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sexo Feminino.	146
8.105	Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sem Distinção de Sexo.	147
8.106	Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa - Sexo Masculino.	148
8.107	Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa entre os 25 e 65 anos - Sexo Masculino.	149
8.108	Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa - Sexo Feminino.	149
8.109	Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa entre os 25 e 65 anos - Sexo Feminino.	149
8.110	Comparação da frequência relativa dos dados da exposição ao risco da população portuguesa e da população da companhia, no sexo masculino (à esquerda) e no sexo feminino (à direita).	150
A.1	Número de Mortes nos Anos 1940-1960, para o Sexo Masculino.	158
A.2	Número de Mortes nos Anos 1961-1980, para o Sexo Masculino.	159
A.3	Número de Mortes nos Anos 1981-2000, para o Sexo Masculino.	159
A.4	Número de Mortes nos Anos 2001-2020, para o Sexo Masculino.	160
A.5	Número de Expostos ao Risco nos Anos 1940-1960, para o Sexo Masculino.	160
A.6	Número de Expostos ao Risco nos Anos 1961-1980, para o Sexo Masculino.	161
A.7	Número de Expostos ao Risco nos Anos 1981-2000, para o Sexo Masculino.	161

A.8	Número de Expostos ao Risco nos Anos 2001-2020, para o Sexo Masculino.	162
A.9	Número de Mortes nos Anos 1940-1960, para o Sexo Feminino.	162
A.10	Número de Mortes nos Anos 1961-1980, para o Sexo Feminino.	163
A.11	Número de Mortes nos Anos 1981-2000, para o Sexo Feminino.	163
A.12	Número de Mortes nos Anos 2001-2020, para o Sexo Feminino.	164
A.13	Número de Expostos ao Risco nos Anos 1940-1960, para o Sexo Feminino.	164
A.14	Número de Expostos ao Risco nos Anos 1961-1980, para o Sexo Feminino.	165
A.15	Número de Expostos ao Risco nos Anos 1981-2000, para o Sexo Feminino.	165
A.16	Número de Expostos ao Risco nos Anos 2001-2020, para o Sexo Feminino.	166
A.17	Número de Mortes nos Anos 1940-1960, Sem Distinção de Sexo.	166
A.18	Número de Mortes nos Anos 1961-1980, Sem Distinção de Sexo.	167
A.19	Número de Mortes nos Anos 1981-2000, Sem Distinção de Sexo.	167
A.20	Número de Mortes nos Anos 2001-2020, Sem Distinção de Sexo.	168
A.21	Número de Expostos ao Risco nos Anos 1940-1960, Sem Distinção de Sexo.	168
A.22	Número de Expostos ao Risco nos Anos 1961-1980, Sem Distinção de Sexo.	169
A.23	Número de Expostos ao Risco nos Anos 1981-2000, Sem Distinção de Sexo.	169
A.24	Número de Expostos ao Risco nos Anos 2001-2020, Sem Distinção de Sexo.	170

Índice de Tabelas

4.1	Transformações usuais de Box-Cox.	31
5.1	Padrões teóricos das FAC e FACP dos modelos de previsão em séries temporais.	48
5.2	Classificação dos métodos de alisamento exponencial.	54
5.3	Notações Utilizadas nos Métodos de Alisamento Exponencial.	55
6.4	Identificação das ordens do modelo ARIMA - Sexo Masculino.	85
6.5	Identificação das ordens do modelo ARIMA - Sexo Feminino.	85
6.6	Identificação das ordens do modelo ARIMA - Sem Distinção de Sexo.	86
6.7	Modelos ARIMA mais adequados para cada caso.	86
6.8	Identificação das ordens do modelo ARIMA - Sexo Feminino.	90
6.9	Identificação das ordens do modelo ARIMA - Sem Distinção de Sexo.	91
6.10	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.	94
6.11	Estimativas iniciais para o nível e declive e estimativas das constantes de alisamento, resultantes da aplicação do Método de Holt à série temporal do nível geral de mortalidade k_t produzido a partir do Método Lee-Carter.	94
6.12	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.	96
6.13	Identificação da ordem do modelo ARIMA - Sexo Masculino.	107
6.14	Identificação da ordem do modelo ARIMA - Sexo Feminino.	107
6.15	Identificação da ordem do modelo ARIMA - Sem Distinção de Sexo.	108
6.16	Modelos ARIMA mais adequados para cada caso.	108

6.17	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA.	112
6.18	Estimativas iniciais para o nível e declive e estimativas das constantes de alisamento, resultantes da aplicação do Método de Holt ao parâmetro k_t produzido a partir da Extensão do Método Lee-Carter.	113
6.19	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.	114
6.20	Modelos ARIMA mais adequados para cada método - Sexo Masculino.	115
6.21	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Masculino.	115
6.22	Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Masculino.	116
6.23	Modelos ARIMA mais adequados para cada método - Sexo Feminino.	120
6.24	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Feminino.	120
6.25	Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Feminino.	121
6.26	Modelos ARIMA mais adequados para cada método - Sem Distinção de Sexo.	123
6.27	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sem Distinção de Sexo.	123
6.28	Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sem Distinção de Sexo.	124

6.29	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Masculino.	126
6.30	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Feminino.	127
6.31	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Feminino.	127
6.32	Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA e Método de Holt.	128
7.33	Cálculo de uma pensão vitalícia.	134
8.34	Equações das Médias Móveis e Médias Móveis Ponderadas.	144
A.1	Medidas descritivas dos dados de exposição ao risco dos três casos de estudo.	170
A.2	Medidas descritivas dos dados de número de mortes dos três casos de estudo.	171
A.3	Médias anuais (de 10 em 10 anos) dos dados de exposição ao risco dos três casos de estudo.	171
A.4	Médias anuais (de 10 em 10 anos) dos dados de número de mortes dos três casos de estudo.	172
B.1	Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sexo Masculino).	173
B.2	Estimação do Parâmetro k_t do Método Lee-Carter (Sexo Masculino).	177
B.3	Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sexo Feminino).	180
B.4	Estimação do Parâmetro k_t do Método Lee-Carter (Sexo Feminino).	184
B.5	Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sem Distinção de Sexo).	187
B.6	Estimação do Parâmetro k_t do Método Lee-Carter (Sem Distinção de Sexo).	190
C.1	Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sexo Masculino).	194
C.2	Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sexo Masculino).	198
C.3	Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sexo Feminino).	201

C.4	Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sexo Feminino).	205
C.5	Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sem Distinção de Sexo).	208
C.6	Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sem Distinção de Sexo).	211
D.1	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Masculino.	215
D.2	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Masculino.	219
D.3	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Feminino.	222
D.4	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Feminino.	226
D.5	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sem Distinção de Sexo.	230
D.6	Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sem Distinção de Sexo.	234
E.1	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Masculino.	238
E.2	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Masculino.	242
E.3	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Feminino.	245
E.4	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Feminino.	249
E.5	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sem Distinção de Sexo.	253

E.6	Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sem Distinção de Sexo.	257
F.1	Medidas descritivas dos dados de exposição ao risco dos três casos de estudo.	261
F.2	Medidas descritivas dos dados de número de mortes dos três casos de estudo.	262
F.3	Médias anuais dos dados de exposição ao risco dos três casos de estudo.	262
F.4	Médias anuais dos dados de número de mortes dos três casos de estudo.	263
G.1	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Masculino.	264
G.2	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Masculino.	265
G.3	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Masculino.	266
G.4	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Masculino.	267
G.5	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Feminino.	268

G.6	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Feminino.	269
G.7	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Feminino.	270
G.8	Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Feminino.	271

Abreviaturas

ADF – *Augmented Dickey-Fuller* (em português, Dickey-Fuller Aumentado)

AIC – *Akaike Information Criterion* (em português, Critério de Informação de Akaike)

AR – *Autoregressive* (em português, Autorregressivo)

ARIMA – *Autoregressive Integrated Moving Average* (em português, Autorregressivo e de Médias Móveis Integrado)

ARMA – *Autoregressive Moving Average* (em português, Autorregressivo e de Médias Móveis)

ASF – Autoridade de Supervisão de Seguros e Fundos de Pensões

BIC – *Bayesian Information Criterion* (em português, Critério de Informação Bayesiano)

DF – Dickey-Fuller

EUA – Estados Unidos da América

EAM – Erro Absoluto Médio

EPAM – Erro Percentual Absoluto Médio

EQM – Erro Quadrático Médio

FAC – Função de Autocorrelação

FACP – Função de Autocorrelação Parcial

HMD – *Human Mortality Database*

INE – Instituto Nacional de Estatística

IPSS – Instituição Particular de Solidariedade Social

KPSS – Kwiatkowski–Phillips–Schmidt–Shin

LC – Lee-Carter

MA – *Moving Average* (em português, Médias Móveis)

MAPE – *Mean Absolute Percentage Error* (em português, Erro Percentual Absoluto Médio)

MSE – *Mean Square Error* (em português, Erro Quadrático Médio)

MV – Máxima Verosimilhança

NA - *Not Available*

PP – Phillips-Perron

REQM - Raiz do Erro Quadrático Médio

RMSE – *Root Mean Square Error* (em português, Raiz do Erro Quadrático Médio)

SARIMA – *Seasonal Autorregressive Moving Average* (em português, Autorregressivo e de Médias Móveis Integrado Sazonal)

SES – *Simple Exponential Smoothing* (em português, Alisamento Exponencial Simples)

SVD – *Singular Value Decomposition* (em português, Decomposição em Valores Singulares)

TE – Taxa de Encargos

1 Introdução

Esta dissertação é realizada no âmbito do Mestrado em Estatística para Ciência de Dados da Universidade do Minho, no contexto de um estágio curricular realizado na companhia de Seguros Ageas.

No espaço de uma década, em Portugal, verificou-se o aumento de 1,77 anos de vida para o total da população, nomeadamente, 1,90 anos para os homens e 1,48 anos para as mulheres. No triénio 2018-2020, a esperança de vida à nascença foi estimada em 81,06 anos, sendo 78,07 anos para os homens e 83,67 anos para as mulheres (INE, 2021). O aumento da esperança de vida tem um impacto significativo na economia e, em particular, no setor segurador com consequências visíveis nas responsabilidades assumidas pelas empresas de seguros.

A Autoridade de Supervisão de Seguros e Fundos de Pensões (ASF) é a autoridade nacional responsável pela regulação e supervisão, quer prudencial quer comportamental, da atividade seguradora, resseguradora, de fundos de pensões e da mediação de seguros. A ASF tem por missão garantir o bom funcionamento do mercado segurador e fundos de pensões em Portugal, de forma a contribuir para a garantia da proteção dos tomadores de seguros, segurados, subscritores, participantes, beneficiários e lesados (ASF, 2015). Desde 1913 que está fixada em lei, em Portugal, a obrigatoriedade das entidades empregadoras repararem as consequências dos acidentes de trabalho sofridos pelos seus trabalhadores. Mais tarde foi instituída a obrigatoriedade legal do seguro pelo risco de acidentes de trabalho, visando assegurar aos trabalhadores por conta de outrem e aos seus familiares condições adequadas de reparação dos danos decorrentes de acidentes de trabalho (ASF, 2015).

O estudo realizado neste trabalho insere-se no ramo de acidente de trabalho (seguros não-vida) e pretende-se, através dos dados da população portuguesa e com os dados da empresa, construir tábuas de mortalidade geracionais que modelem o comportamento da população portuguesa e dos pensionistas da companhia. Tradicionalmente, os atuários baseiam-se em tábuas contemporâneas (ou de momento) para o cálculo dos prémios de seguro, sendo que estas são construídas a partir de quocientes de mortalidade estimados com base nas estimativas demográficas observadas num

determinado momento. Apesar desta abordagem ser simples e pragmática, esta revela-se claramente desajustada no caso em que a mortalidade evolui ao longo do tempo. Na prática, os prémios calculados a partir destes dados não refletem o risco real dos pensionistas pelo que, em contexto de aumento da longevidade, a realização do *pricing* nestes termos leva a uma subestimação dos prémios. Em Portugal, a construção de tábuas de mortalidade geracionais ocorreu somente a partir do início da década de 2000, pelo Instituto Nacional de Estatística, sendo que antes disso as companhias de seguros eram "forçadas" a recorrerem a tábuas de mortalidade adotadas noutros países como, por exemplo, em França e Espanha. Note-se que, embora se trate de uma prática permitida pelas autoridades de supervisão, a utilização de uma lei de sobrevivência relativa à população de um outro país, tendencialmente desajustada das condições demográficas observadas no país em que é aplicada, envolve naturalmente o risco de sobrestimação da taxa de mortalidade da população. Torna-se por isso indispensável a construção de tábuas de mortalidade regulamentares, quer para a população em geral, quer para as subpopulações de pessoas seguras e dos fundos de pensões.

1.1 Companhia de Seguros Ageas

A companhia de seguros Ageas é uma empresa multinacional com representação em diversos países como a Bélgica, o Reino Unido, a França, a Turquia, Portugal e, ainda, nove países da Ásia, incluindo a China, a Malásia, a Tailândia, a Índia, Filipinas e Vietnam. O principal objetivo da empresa é proporcionar tranquilidade aos clientes quando estes mais necessitam e o seu papel, como seguradora, incide em ajudar os clientes em todas as fases da sua vida a mitigar riscos relacionados a bens, sinistros, vida e pensões.

A operar em Portugal desde 2005, a Ageas aposta no país como um dos principais mercados onde se pretende desenvolver, através de parcerias fortes, e contribuindo para o desenvolvimento do país e da sociedade através da Fundação Ageas (Instituição Particular de Solidariedade Social (IPSS) fundada em 1998, agregadora de pessoas e parceiros, que visa a prossecução de fins de solidariedade social na comunidade), ajudando os clientes a gerir, antecipar e proteger-se contra riscos e imprevistos, para que possam viver o presente e o futuro com a máxima segurança e serenidade.

O grupo Ageas Portugal é constituído por várias marcas comerciais: Ageas Seguros, Ageas Pensões, Médis, Ocidental e Seguro Direto, sendo um dos líderes no *ranking* segurador português. A Ageas Seguros disponibiliza produtos a particulares nos ramos de saúde, acidentes pessoais, automóvel, viagens e lazer,

casa, acidentes de trabalho, casa e vida (dependem diretamente da vida ou morte da pessoa segura).

1.1.1 Ramo Acidentes de Trabalho

Um acidente de trabalho é aquele que se verifica no local e no tempo de trabalho e que produza direta ou indiretamente lesão corporal, perturbação funcional ou doença de que resulte redução na capacidade de trabalho, ou de ganho, ou a morte (ASF, 2015). Também são considerados acidentes de trabalho: o trajeto de ida e volta para o trabalho; entre o local de trabalho e o espaço de refeição; entre o local que o colaborador possa estar a trabalhar e o regresso para o seu trabalho habitual ou casa.

Como foi referido anteriormente, esta dissertação enquadra-se no âmbito do ramo de acidentes de trabalho. Neste âmbito, a companhia Ageas Seguros oferece seguros para trabalhadores independentes, que trabalham por conta própria e para empresas, de forma a proteger os empregados domésticos. Estes seguros são de cariz fundamental, visto que são seguros de caráter obrigatório para todos os trabalhadores, como definido pela Constituição da República Portuguesa na Lei n.º 100/97, de 13 de setembro.

Relativamente aos seguros para empresas, a companhia garante o pagamento de baixas, pensões e subsídios, assistência médica e cirúrgica, assistência medicamentosa e farmacêutica, cuidados de enfermagem, hospitalização e serviço de reabilitação.

1.2 Objetivos do Trabalho

O principal objetivo desta dissertação é, como foi referido anteriormente, construir tábuas de mortalidade geracionais que modelem o comportamento da população portuguesa e dos pensionistas da companhia, através dos dados da população portuguesa e dos dados da empresa.

Nesta dissertação pretende-se desenvolver modelos que permitam estimar os fatores a serem usados para o cálculo das pensões vitalícias no ramo de Acidentes de Trabalho, sendo que, inicialmente, são modelados os dados referentes à população portuguesa e, em seguida, são efetuados os ajustamentos aos dados específicos da empresa.

De realçar que toda a análise estatística efetuada nesta dissertação foi realizada através do *software R* e do *Microsoft Excel*.

1.3 Estrutura do Trabalho

Nesta secção é descrita a estrutura do trabalho. Esta dissertação encontra-se dividida em nove capítulos.

No Capítulo 1 é realizado um breve enquadramento do tema do estudo, assim como uma breve apresentação da empresa e a descrição dos principais objetivos que se pretendem alcançar.

No Capítulo 2 são descritos os tipos de tábuas de mortalidade existentes e as funções que as constituem, bem como são calculadas as respetivas funções.

No Capítulo 3 descrevem-se alguns métodos de ajustamento da taxa de mortalidade, nomeadamente, o Método Lee-Carter e o Modelo de Poisson. Neste Capítulo são apresentadas as equações de cada método, bem como os parâmetros dos modelos e os métodos de estimação e modelação utilizados em cada um dos casos. No final do capítulo são também apresentadas algumas limitações observadas na utilização desses métodos.

Nos Capítulos 4 e 5 são descritos os conteúdos teóricos relacionados com os métodos de previsão de séries temporais, mais concretamente os conceitos fundamentais sobre séries temporais, os métodos que são aplicados aos dados e as medidas de avaliação utilizadas para selecionar o método mais adequado ao estudo. A aplicação das metodologias descritas são apresentadas nos Capítulos 6 e 7. No Capítulo 6, inicialmente, é elaborada uma breve análise exploratória dos dados da população portuguesa, com e sem distinção de sexo, seguindo-se a aplicação do Método Lee-Carter e do Modelo de Poisson. Por fim, é realizado um estudo comparativo da capacidade preditiva dos diferentes métodos utilizados, através das medidas de avaliação apresentadas. O Capítulo 7 também é referente aos dados da população portuguesa, sendo que neste são apresentadas algumas funções mais relevantes das tábuas de mortalidade e a sua evolução ao longo do tempo, e também são calculadas as taxas vitalícias que são necessárias para que, posteriormente, sejam calculadas as pensões vitalícias dos pensionistas.

O Capítulo 8 refere-se à aplicação das metodologias aos dados da companhia, sendo que inicialmente é elaborada uma análise exploratória dos dados que, tal como no caso da população portuguesa, é realizada com e sem distinção de sexo. De seguida, são calculadas as taxas de mortalidade de cada ano e estas são comparadas com as taxas obtidas para a população portuguesa.

Por fim, no Capítulo 9 são apresentadas as principais conclusões do trabalho desenvolvido e sobre os resultados obtidos, assim como algumas sugestões para trabalho futuro.

Capítulo 2

Tábuas de Mortalidade

A atividade seguradora é uma atividade indispensável para o correto funcionamento da sociedade, potencia a iniciativa empreendedora e a criação de valor, protegendo as famílias e as empresas face aos riscos, salvaguardando e estabilizando a sua situação financeira (Batista, 2021).

Para ser possível a realização do seguro é necessário entender o comportamento da população/pensionistas da companhia, através da esperança média de vida, por sexo e idades, o nível de mortalidade e longevidade de uma população, entre outros. Para esta finalidade são utilizadas as tábuas de mortalidade, que são uma ferramenta estatística utilizada por diversos especialistas como demógrafos, atuários, entre outros.

As tábuas de mortalidade são um instrumento indispensável para a ciência atuarial, uma vez que a partir destas é possível modelar o risco de morte de uma população num exato momento. Uma tábua de mortalidade é um modelo estatístico que combina as taxas de mortalidade às diferentes idades, transformando-as em quocientes de mortalidade. Através da associação de um conjunto de funções básicas permite medir o fenómeno de mortalidade, deduzir as probabilidades de sobrevivência e a esperança média de vida (Carrilho & Patrício, 2004). Estas têm um relevante interesse no ramo atuarial uma vez que complementam os cálculos relativos às operações dos seguros e são fundamentais para a elaboração de projeções demográficas para a população.

Considera-se que as primeiras tábuas de mortalidade surgiram na época dos Césares, na antiga Roma, com a chamada Tábua Ulpiana, datada do ano 220 (Bravo, 2007). Devido à falta de informação existente na altura, estas eram consideradas rudimentares e não indicavam o verdadeiro comportamento da mortalidade. A partir do momento em que se iniciou o processo de recolha de dados sobre a mortalidade das populações, já foi possível construir tábuas de mortalidade mais rigorosas.

Deste modo, com base no número de óbitos registados nas paróquias de Londres, John Graunt construiu a primeira tábua de mortalidade. A partir deste momento, e devido a algumas fragilidades metodológicas existentes nas tábuas construídas por Graunt, foram várias as investigações elaboradas em vários domínios da Demografia, como por exemplo por Jan de Witt (1671), primeiro-ministro holandês, Edmund Halley (1693), entre outros. Apesar de todas as investigações e criações de tábuas de mortalidade realizadas após Graunt, somente em 1815 é que é elaborada a primeira tábua considerada cientificamente correta, esta era apoiada em população e óbitos classificados por idade, referentes a duas paróquias da cidade britânica de Carlisle, no período 1779-1787. Desde então foram divulgadas várias tábuas em vários países.

Em Portugal, o INE publica tábuas de mortalidade desde 1930, porém nesta altura eram construídas tábuas contemporâneas em vez de tábuas geracionais. A construção das tábuas de mortalidade e o cálculo da esperança média de vida da população é um estudo estatístico que se realiza anualmente, abrangendo toda a população de Portugal, com base em informação proveniente de outras operações estatísticas, da área da Demografia, desenvolvidas pelo INE (Coelho & Malta, 2014). Através de um estudo realizado pelo INE, prevê-se que o índice de envelhecimento em Portugal quase duplicará, passando de 159 para 300 idosos por cada 100 jovens, em 2080, em resultado do decréscimo da população jovem e do aumento da população idosa (INE, 2020).

2.1 Classificação das Tábuas de Mortalidade

As tábuas de mortalidade podem ser classificadas quanto às gerações utilizadas nas mesmas:

- **Tábuas Contemporâneas (Transversais ou de Momento):**

Baseiam-se na análise *cross-section* de uma geração fictícia (normalmente composta por 100 000 indivíduos) que é sujeita às taxas específicas de mortalidade observadas para cada idade (ou grupos de idade) num dado ano (ou num período de tempo). Para a construção destas tábuas são normalmente utilizadas estatísticas referentes a um curto período de tempo, no caso do número de óbitos e da população exposta ao risco.

- **Tábuas Longitudinais (ou Geracionais):**

Baseiam-se no cálculo das taxas de mortalidade registadas por uma mesma geração, isto é, indivíduos nascidos no mesmo ano. A realização deste tipo de tábuas exige que as pessoas

sejam seguidas desde o momento do seu nascimento até à sua morte, ou seja, ao longo de toda a sua vida. Este procedimento designa-se por análise longitudinal da mortalidade.

As tábuas de mortalidade podem também ser classificadas quanto à amplitude do intervalo de idades no qual assenta a informação:

- **Tábuas Completas:**

Contêm dados para cada idade singular, isto é, idade ano a ano, desde o nascimento até ao limite superior adotado;

- **Tábuas Abreviadas:**

Contêm dados para grupos de idades quinquenais ou decenais.

Outra classificação possível passa pela análise da idade em que o indivíduo passa a integrar a população, ou seja, a idade com que este é selecionado. A partir desta é possível construir tábuas bidimensionais, em que os índices são a idade de seleção e a duração desde esse momento. Neste caso podem ser distinguidas as seguintes tábuas:

- **Tábuas Selecionadas:**

As probabilidades são tabuladas considerando a idade x do indivíduo e outras informações, por exemplo, o tempo decorrido desde a realização de uma apólice de seguro;

- **Tábuas Finais:**

Correspondem à última coluna de uma tábua selecionada, ou seja, aquelas em que se admite que o período de seleção deixa de ter efeito na mortalidade.

Por fim, as tábuas de mortalidade podem ser classificadas consoante o tempo cronológico utilizado na elaboração e apresentação das mesmas:

- **Tábuas Estáticas:**

São tábuas unidimensionais onde todas as funções são atribuídas apenas à idade biológica x ;

- **Tábuas Dinâmicas (ou Prospetivas):**

São tábuas bidimensionais em que todas as funções da mesma são indexadas, em linha, pela idade biológica x e, em coluna, pelo ano cronológico t .

Neste trabalho são elaboradas tábuas geracionais, completas e dinâmicas, uma vez que para cada tábua é utilizada a mesma geração, ou seja, indivíduos que nascem no mesmo ano e estas são calculadas para cada idade singular, isto é, desde o nascimento até ao limite superior adotado, e para cada ano de calendário.

2.2 Projeção de Tábuas de Mortalidade

Uma tábua de mortalidade é um elemento de análise demográfica que sintetiza um certo número de funções que permitem analisar, numa dada população, o fenómeno de mortalidade e deduzir a vida média dessa população (Carrilho & Patrício, 2004).

Uma tábua de mortalidade é constituída por seis funções básicas, que são: q_x , l_x , d_x , L_x , T_x e e_x . As três primeiras funções são consideradas medidas de frequência ou de intensidade dos acontecimentos, já as restantes são medidas de duração de tempo de vida de indivíduos, em anos.

Para ser possível realizar o cálculo de todas as funções, inicialmente, é necessário definir e calcular a taxa de mortalidade. A taxa de mortalidade de um indivíduo para a idade x no ano t , $m_{x,t}$, é definida como a relação entre o número de mortes $d_{x,t}$, registadas durante o ano civil t para pessoas de idade x e o número de pessoas expostas ao risco com idade x no ano t , $E_{x,t}$ (Russolillo, 2014),

$$m_{x,t} = \frac{d_{x,t}}{E_{x,t}}. \quad (2.1)$$

No caso deste trabalho o número de pessoas expostas ao risco constitui o número de pessoas a quem está a ser paga uma pensão.

Após serem calculadas as taxas de mortalidade, é possível calcular todas as funções básicas que integram uma tábua de mortalidade:

- q_x : Quociente de mortalidade - probabilidade de um indivíduo com idade x falecer antes de atingir a idade $x + 1$

$$q_x = \frac{w_x m_x}{(1 + f'_x w_x m_x)}, \quad (2.2)$$

onde m_x representa a taxa de mortalidade, w_x indica o número de idades presentes dentro de um grupo (neste estudo serão consideradas idades isoladas, logo w_x será igual a 1) e f'_x é o número médio de anos vividos dentro do intervalo $[x, x + 1]$ para pessoas que morrem com essa idade (Russolillo, 2014) (sendo que $f'_x = 1 - f_x$ e assume-se que $f_x = \frac{1}{2}$ para todos os grupos exceto quando $x = 0$, nesse caso $f_x = 0,15$ para homens e $f_x = 0,16$ para mulheres);

- l_x : Número de sobreviventes à idade exata x - número de indivíduos que sujeitos à mortalidade de q_x atingem a idade x , considerando uma geração inicial de 100 000 nados-vivos

$$l_{x+1} = l_x(1 - q_x); \quad (2.3)$$

- d_x : Número de óbitos entre as idades exatas x e $x + 1$ - número de óbitos da geração inicial registados entre as idades x e $x + 1$

$$d_x = l_x - l_{x+1} = l_x q_x; \quad (2.4)$$

- L_x : Sobreviventes, em anos completos, entre as idades exatas x e $x + 1$ - total de anos completos vividos pelos sobreviventes da geração inicial entre as idades exatas x e $x + 1$

$$L_x = \frac{d_x}{m_x}; \quad (2.5)$$

- T_x : Anos completos após a idade x - total de anos completos vividos pelos sobreviventes após a idade x

$$T_x = \sum_{x_i}^{x_n} L_{x_i}; \quad (2.6)$$

- e_x : Esperança de vida à idade x - número esperado de anos de vida futura dos indivíduos que atingiram a idade x

$$e_x = \frac{T_x}{l_x}. \quad (2.7)$$

A título de exemplo, na Figura 2.1 está disponível um excerto da tábua de mortalidade Portuguesa de 2018-2020, para ambos os sexos, realizada pelo INE.

Tábua Completa de Mortalidade para Portugal 2018-2020 (Ambos os sexos)						
Idade	Quociente de mortalidade	Sobreviventes à idade exata x	Óbitos entre as idades exatas x e x+1	Sobreviventes entre as idades exatas x e x+1	Anos completos após a idade x	Esperança de vida
(x)	(qx)	(lx)	(dx)	(Lx)	(Tx)	(ex)
1	2	3	4	5	6	7
0	0,002972	100 000	297	99 813	8 106 311	81,06
1	0,000316	99 703	32	99 687	8 006 498	80,30
2	0,000172	99 671	17	99 663	7 906 811	79,33
3	0,000120	99 654	12	99 648	7 807 148	78,34
4	0,000106	99 642	11	99 637	7 707 500	77,35
5	0,000078	99 632	8	99 628	7 607 863	76,36
6	0,000057	99 624	6	99 621	7 508 236	75,37
7	0,000064	99 618	6	99 615	7 408 615	74,37
8	0,000071	99 612	7	99 608	7 309 000	73,37
9	0,000096	99 605	10	99 600	7 209 392	72,38
10	0,000091	99 595	9	99 590	7 109 792	71,39
11	0,000065	99 586	6	99 583	7 010 201	70,39
12	0,000075	99 579	7	99 576	6 910 619	69,40
13	0,000101	99 572	10	99 567	6 811 043	68,40
14	0,000090	99 562	9	99 557	6 711 476	67,41
15	0,000156	99 553	16	99 545	6 611 918	66,42
16	0,000200	99 537	20	99 527	6 512 373	65,43
17	0,000172	99 518	17	99 509	6 412 846	64,44
18	0,000249	99 500	25	99 488	6 313 337	63,45
19	0,000229	99 476	23	99 464	6 213 849	62,47
20	0,000302	99 453	30	99 438	6 114 384	61,48
21	0,000317	99 423	32	99 407	6 014 947	60,50
22	0,000354	99 391	35	99 374	5 915 540	59,52
23	0,000355	99 356	35	99 338	5 816 166	58,54

Figura 2.1: Excerto da tábua de mortalidade Portuguesa de 2018-2020, para ambos os sexos, realizada pelo Instituto Nacional de Estatística (INE).

Capítulo 3

Método Lee-Carter e Modelo de Poisson (Extensão do Método Lee-Carter)

Neste capítulo são apresentados os dois métodos utilizados para obter as previsões das taxas de mortalidade, $m_{x,t}$. Primeiramente, é definido o Método Lee-Carter e, posteriormente, o Modelo de Poisson que também pode ser denominado como a Extensão do Método Lee-Carter.

3.1 Método Lee-Carter

O método Lee-Carter (LC) foi um método pioneiro para o cálculo de previsões das taxas de mortalidade específicas por grupo etário e, conseqüentemente, para a esperança de vida (Coelho, 2005). Este é um método designado para a extrapolação das tendências e dos padrões etários da mortalidade específicas por idade e grupo etário. Este método foi inicialmente aplicado a dados de mortalidade da população dos EUA no período 1900-1989 com um horizonte de previsão de 76 anos e, devido à qualidade dos resultados obtidos nas várias aplicações, tem vindo a ser aplicado em diversos países. Em Portugal este método também tem sido utilizado, sendo que as últimas projeções publicadas pelo INE foram realizadas para o horizonte de 2080. Lee e Carter (R. D. Lee & Carter, 1992) desenvolveram um método que define a decomposição da mortalidade em duas parcelas: um modelo demográfico para a mortalidade, específico para cada idade, e um indicador de tendência temporal, ao qual são aplicados métodos de séries temporais de maneira a projetar valores futuros para os indicadores de mortalidade. O modelo demográfico é estimado pela informação já existente da mortalidade, onde se obtêm dois índices referentes à idade e um índice referente ao ano. De seguida, o

índice referente ao ano é modelado como uma série temporal e, posteriormente, projetado, obtendo-se valores futuros de taxas de mortalidade.

Os autores destacam como virtudes deste novo método o facto de combinar um modelo demográfico rico, mas parcimonioso, com um modelo de séries temporais, permitindo obter intervalos probabilísticos para as respetivas previsões, e de se basear num conjunto de informação histórica num horizonte temporal relativamente longo. Contudo, referem que se trata de um método extrapolativo, ou seja, que se baseia no comportamento da mortalidade no passado para prever o futuro, e embora tenha algumas vantagens sobre outros métodos extrapolativos, partilha as fraquezas fundamentais da extrapolação. A estrutura e a evolução verificada nos dados históricos poderão não se manter no futuro, podendo existir alterações estruturais que em termos de previsão não irão ser contempladas (Coelho, 2005).

Antes de ser explicado o Método Lee-Carter, é necessário rever a definição da taxa de mortalidade. Como foi explicado no Capítulo 2, a taxa de mortalidade é o quociente entre o número de mortes, $d_{x,t}$, e o número de expostos ao risco, $E_{x,t}$, e indica o número de óbitos por 1000 habitantes de uma população num determinado intervalo de tempo, normalmente, um ano. Esta apresenta um comportamento inverso à esperança média de vida no momento de nascimento, uma vez que, quanto maior é a esperança de vida, menor é a taxa de mortalidade. A taxa de mortalidade habitualmente considera-se elevada quando excede os 30%, moderada quando se situa entre os 15% e os 30% e baixa quando fica abaixo dos 15% (Ferreira, 2012). Como era de esperar, esta taxa é mais elevada em países não desenvolvidos do que em países desenvolvidos.

O método LC descreve a evolução da mortalidade e é definido pela representação log-bilinear da taxa de mortalidade

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}, \quad (3.1)$$

onde

- α_x - indica o nível médio das taxas de mortalidade no tempo, para cada idade, isto é, descreve a forma geral do perfil de mortalidade por idade;
- β_x - representa as oscilações nas taxas de mortalidade na idade x , em resposta a alterações no nível geral de mortalidade k_t . Ou seja, permite averiguar a sensibilidade de $m_{x,t}$ às variações de k_t ;
- k_t - constitui o índice de nível geral da mortalidade no ano t , isto é, identifica a tendência temporal principal do nível de mortalidade;

- $\epsilon_{x,t}$ - representa o erro aleatório do tipo ruído branco com média zero e variância σ_ϵ^2 , reflete as influências históricas específicas a cada idade não capturadas pelo modelo.

3.2 Extensão do Método Lee-Carter (Modelo de Poisson)

Uma das principais críticas apontadas ao Método Lee-Carter refere-se à hipótese de que os erros são homocedásticos, ou seja, que apresentam variância constante. Esta suposição é bastante irrealista uma vez que, devido à existência de um número reduzido de óbitos em idades avançadas, o logaritmo da taxa de mortalidade é muito mais variável nestas idades do que nas idades jovens (Bravo, 2007).

De forma a contornar este problema, (Brouhns et al., 2002) propuseram uma extensão do Método LC, considerando que o número de óbitos segue uma distribuição de Poisson

$$D_{x,t} \sim \text{Poisson}(\mu_{x,t}E_{x,t}), \quad (3.2)$$

com

$$\mu_{x,t} = \exp(\alpha_x + \beta_x k_t), \quad (3.3)$$

onde os parâmetros α_x , β_x e k_t mantêm, na sua essência, o significado originalmente atribuído pelo Método Lee-Carter.

3.3 Ajustamento do Modelo Demográfico

3.3.1 Método Lee-Carter

Para o ajustamento do modelo demográfico, Lee e Carter decidiram realizar uma aproximação ao Método de Decomposição em Valores Singulares (SVD), procedendo a este ajustamento em duas fases. Inicialmente serão estimados os parâmetros α_x , β_x e k_t e, posteriormente, será realizada uma reestimação do parâmetro k_t , para que o número de óbitos ajustado seja igual ao número de óbitos efetivamente verificado.

A estimação dos parâmetros α_x , β_x e k_t é efetuada pelo método dos mínimos quadrados, com a decomposição em valores singulares (SVD). O método SVD é aplicado à matriz dos logaritmos das taxas depois de lhes ter sido subtraída a média temporal do logaritmo das taxas de mortalidade específicas por idade.

O modelo 3.1 é sobreparametrizado, ou seja, o modelo é invariante a algumas transformações. Por exemplo, dada uma constante c

$$\{\alpha_x, \beta_x, k_t\} \mapsto \{\alpha_x, \beta_x/c, ck_t\} \forall c \in \mathbb{R}, c \neq 0 \quad (3.4)$$

$$\{\alpha_x, \beta_x, k_t\} \mapsto \{\alpha_x - c\beta_x, \beta_x, k_t + c\} \forall c \in \mathbb{R}. \quad (3.5)$$

Uma vez que o sistema definido em 3.1 é indeterminado, é necessário impor condições adicionais. Dessa forma, os autores do método LC adicionaram duas restrições ao modelo de modo a obter apenas uma única solução para a estimação de cada parâmetro. Essas restrições são

$$\sum_{x=x_{min}}^{x_{max}} \beta_x = 1, \quad (3.6)$$

$$\sum_{t=t_{min}}^{t_{max}} k_t = 0, \quad (3.7)$$

onde x_{min} e x_{max} representam, respectivamente, a idade inicial e a idade final e t_{min} e t_{max} traduzem, respectivamente, o ano inicial e o ano final.

Sob estes pressupostos é possível realizar a estimação de cada parâmetro como

- **Estimação de α_x**

$$\begin{aligned} & \sum_{x,t} ((\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)^2 = 0 \\ \Leftrightarrow & \sum_{t=t_{min}}^{t_{max}} \ln(m_{x,t}) = \sum_{t=t_{min}}^{t_{max}} (\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t) \\ \Leftrightarrow & \sum_{t=t_{min}}^{t_{max}} \hat{\alpha}_x = \sum_{t=t_{min}}^{t_{max}} \ln(m_{x,t}) - \sum_{t=t_{min}}^{t_{max}} \hat{\beta}_x \hat{k}_t \\ \Leftrightarrow & \hat{\alpha}_x \sum_{t=t_{min}}^{t_{max}} 1 = \sum_{t=t_{min}}^{t_{max}} \ln(m_{x,t}) - \hat{\beta}_x \sum_{t=t_{min}}^{t_{max}} \hat{k}_t \\ \Leftrightarrow & (t_{max} - t_{min} + 1) \hat{\alpha}_x = \sum_{t=t_{min}}^{t_{max}} \ln(m_{x,t}) \\ \Leftrightarrow & \hat{\alpha}_x = \frac{1}{h} \sum_{t=t_{min}}^{t_{max}} \ln(m_{x,t}), \end{aligned} \quad (3.8)$$

onde $h = t_{max} - t_{min} + 1$.

- **Estimação de k_t**

$$\begin{aligned}
& \sum_{x,t} (\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)^2 = 0 \\
\Leftrightarrow & \sum_{x=x_{min}}^{x_{max}} \ln(m_{x,t}) = \sum_{x=x_{min}}^{x_{max}} \hat{\alpha}_x + \sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x \hat{k}_t \\
\Leftrightarrow & \sum_{x=x_{min}}^{x_{max}} (\ln(m_{x,t}) - \hat{\alpha}_x) = \hat{k}_t \sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x \\
\Leftrightarrow & \hat{k}_t = \sum_{x=x_{min}}^{x_{max}} (\ln(m_{x,t}) - \hat{\alpha}_x). \tag{3.9}
\end{aligned}$$

- **Estimação de β_x**

Cada $\hat{\beta}_x$ é estimado através do método dos Mínimos Quadrados, isto é, de forma a minimizar a seguinte expressão

$$\sum_{x,t} ((\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)^2. \tag{3.10}$$

Derivando 3.10 em ordem a $\hat{\beta}_x$,

$$\sum_{t=t_{min}}^{t_{max}} 2((\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)(-\hat{k}_t). \tag{3.11}$$

Para encontrar o mínimo da função é necessário igualar a primeira derivada a zero

$$\begin{aligned}
& 2 \sum_{t=t_{min}}^{t_{max}} ((\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)(-\hat{k}_t) = 0 \\
\Leftrightarrow & -1 \sum_{t=t_{min}}^{t_{max}} ((\ln(m_{x,t}) - \hat{\alpha}_x - \hat{\beta}_x \hat{k}_t)(\hat{k}_t) = 0 \\
\Leftrightarrow & \sum_{t=t_{min}}^{t_{max}} \hat{k}_t (\ln(m_{x,t}) - \hat{\alpha}_x) - \sum_{t=t_{min}}^{t_{max}} \hat{\beta}_x \hat{k}_t^2 = 0 \\
\Leftrightarrow & \hat{\beta}_x \sum_{t=t_{min}}^{t_{max}} \hat{k}_t^2 = \sum_{t=t_{min}}^{t_{max}} \hat{k}_t (\ln(m_{x,t}) - \hat{\alpha}_x) \\
\Leftrightarrow & \hat{\beta}_x = \frac{\sum_{t=t_{min}}^{t_{max}} \hat{k}_t (\ln(m_{x,t}) - \hat{\alpha}_x)}{\sum_{t=t_{min}}^{t_{max}} \hat{k}_t^2}. \tag{3.12}
\end{aligned}$$

e averiguar que a segunda derivada da função é sempre positiva.

Como as estimativas de α_x , β_x e k_t foram calculadas a partir da matriz de logaritmos das taxas de mortalidade e não pela matriz das taxas de mortalidade, o número de óbitos obtidos pela estimação não coincide com o número de óbitos observados, pelo que é necessário reestimar o parâmetro k_t . Sendo assim, Lee e Carter introduziram uma nova etapa para a estimação de k_t , onde o ajustamento deste parâmetro, utilizando $\hat{\alpha}_x$ e $\hat{\beta}_x$ estimados anteriormente, resulte no número total de óbitos observados para o ano em causa. Considerando,

$$\sum_{x=x_{min}}^{x_{max}} d_{x,t} \quad (3.13)$$

como o total de óbitos observado em cada ano e

$$\begin{aligned} \ln(m_{x,t}) &= \hat{\alpha}_x + \hat{\beta}_x \hat{k}_t \\ \Leftrightarrow \ln\left(\frac{d_{x,t}}{E_{x,t}}\right) &= \hat{\alpha}_x + \hat{\beta}_x \hat{k}_t \\ \Leftrightarrow \hat{d}_{x,t} &= \sum_{x=x_{min}}^{x_{max}} E_{x,t} \exp(\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t) \end{aligned} \quad (3.14)$$

o número de mortes estimadas. Os novos estimadores de \hat{k}_t obtêm-se resolvendo a seguinte equação

$$\sum_{x=x_{min}}^{x_{max}} \left[E_{x,t} \exp(\hat{\alpha}_x) \exp(\hat{\beta}_x \hat{k}_t) - d_{x,t} \right] = 0. \quad (3.15)$$

Para resolver este problema aplica-se o método iterativo de Newton-Raphson, tendo em conta os valores brutos de k_t estimados anteriormente. O método Newton-Raphson baseia-se no seguinte esquema iterativo

$$k_t^{(\nu+1)} = k_t^{(\nu)} - \frac{F\left(k_t^{(\nu)}\right)}{F'\left(k_t^{(\nu)}\right)}, \quad (3.16)$$

onde $F(k_t)$ representa a função a anular.

3.3.2 Extensão do Método Lee-Carter (ou Modelo de Poisson)

Este método apresenta uma vantagem importante face ao método LC, uma vez que neste os parâmetros são estimados pelo método de máxima verosimilhança, sem necessidade de recorrer à SVD

ou ao método de mínimos quadrados. Posto isto, designa-se por $\lambda_{x,t}$ o número esperado de óbitos na idade x e no ano t , isto é,

$$\lambda_{x,t} = E[D_{x,t}] = E_{x,t} \exp(\alpha_x + \beta_x k_t) \quad (3.17)$$

e por $d_{x,t}$ o respetivo valor esperado de $D_{x,t}$.

A estimação dos parâmetros será realizada de forma a maximizar a função log-verosimilhança dada por

$$\begin{aligned} L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{k}) &= \ln \left\{ \prod_{t=t_{min}}^{t_{max}} \prod_{x=x_{min}}^{x_{max}} \left(\frac{\lambda_{x,t}^{d_{x,t}} \exp(-\lambda_{x,t})}{(d_{x,t})!} \right) \right\} \\ &= \sum_{t=t_{min}}^{t_{max}} \sum_{x=x_{min}}^{x_{max}} \left\{ d_{x,t} \ln(\lambda_{x,t}) - \lambda_{x,t} - \ln[(d_{x,t})!] \right\} \\ &= \sum_{t=t_{min}}^{t_{max}} \sum_{x=x_{min}}^{x_{max}} \left\{ d_{x,t} \ln(E_{x,t}) + d_{x,t} (\alpha_x + \beta_x k_t) - E_{x,t} \exp(\alpha_x + \beta_x k_t) - \ln[(d_{x,t})!] \right\} \\ &= \sum_{t=t_{min}}^{t_{max}} \sum_{x=x_{min}}^{x_{max}} \left\{ d_{x,t} (\alpha_x + \beta_x k_t) - E_{x,t} \exp(\alpha_x + \beta_x k_t) \right\} + c, \end{aligned} \quad (3.18)$$

onde $\boldsymbol{\alpha} = (\alpha_{x_{min}}, \dots, \alpha_{x_{max}})$, $\boldsymbol{\beta} = (\beta_{x_{min}}, \dots, \beta_{x_{max}})$, $\mathbf{k} = (k_{t_{min}}, \dots, k_{t_{max}})$ e c é uma constante.

Devido à presença do termo bilinear $\beta_x k_t$ não é possível estimar o modelo proposto através da aplicação de um modelo linear generalizado. Para resolver este problema é utilizado o método iterativo de Goodman (baseado no algoritmo de Newton-Raphson) que é um método para estimar modelos log-lineares com termos bilineares, sendo que em cada iteração procura convergir para a estimativa de máxima verosimilhança do parâmetro a estimar. Este método consiste em, na iteração $\nu + 1$, estimar o parâmetro θ , fixando os restantes parâmetros, isto é,

$$\hat{\theta}^{(\nu+1)} = \hat{\theta}^{(\nu)} - \frac{\partial L^{(\nu)} / \partial \theta}{\partial^2 L^{(\nu)} / \partial \theta^2}, \quad (3.19)$$

onde $L^{(\nu)} = L^{(\nu)}(\hat{\theta}^{(\nu)})$.

Antes de se estimarem os parâmetros α_x , β_x e k_t , começa-se por definir os valores iniciais destes parâmetros como $\hat{\alpha}_x^{(0)} = 0$, $\hat{\beta}_x^{(0)} = 1$ e $\hat{k}_t^{(0)} = 0$ (Brouhns et al., 2002) e a atualização destes é realizada da seguinte forma

$$\begin{aligned}
\hat{\alpha}_x^{(\nu+1)} &= \hat{\alpha}_x^{(\nu)} - \frac{\sum_{t=t_{min}}^{t_{max}} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu)})]}{-\sum_{t=t_{min}}^{t_{max}} [E_{x,t} \exp(\hat{\alpha}_x^{(\nu)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu)})]}, \\
\hat{k}_t^{(\nu+1)} &= \hat{k}_t^{(\nu)} - \frac{\sum_{x=x_{min}}^{x_{max}} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu)})] \hat{\beta}_x^{(\nu)}}{-\sum_{x=x_{min}}^{x_{max}} [E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu)})] (\hat{\beta}_x^{(\nu)})^2}, \\
\hat{\beta}_x^{(\nu+1)} &= \hat{\beta}_x^{(\nu)} - \frac{\sum_{t=t_{min}}^{t_{max}} [d_{x,t} - E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu+1)})] \hat{k}_t^{(\nu+1)}}{-\sum_{t=t_{min}}^{t_{max}} [E_{x,t} \exp(\hat{\alpha}_x^{(\nu+1)} + \hat{\beta}_x^{(\nu)} \hat{k}_t^{(\nu+1)})] (\hat{k}_t^{(\nu+1)})^2}. \tag{3.20}
\end{aligned}$$

O critério de paragem consiste num pequeno incremento da função de log-verosimilhança, sendo que o recomendado é da ordem de grandeza de 10^{-10} . Ao contrário do Método LC, quando os parâmetros são estimados pelo método de MV, o número de óbitos observados e estimados são iguais em cada idade e ao longo do período de observação, não sendo necessário reestimar o parâmetro \hat{k}_t . As estimativas de $\hat{\alpha}_x$, $\hat{\beta}_x$ e \hat{k}_t obtidas a partir do método de MV não garantem o cumprimento das restrições 3.6 e 3.7, pelo que é necessário introduzir uma reparametrização

$$k_t^* = (\hat{k}_t - \bar{k}) \sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x \quad e \quad \beta_x^* = \frac{\hat{\beta}_x}{\sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x}, \tag{3.21}$$

onde \bar{k} representa a média aritmética de \hat{k}_t , isto é,

$$\bar{k} = \frac{1}{T} \sum_{t=t_{min}}^{t_{max}} \hat{k}_t. \tag{3.22}$$

É ainda necessário respeitar a igualdade

$$\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t = \alpha_x^* + \beta_x^* k_t^*, \tag{3.23}$$

deduzindo que

$$\begin{aligned}
\alpha_x^* &= \hat{\alpha}_x + \hat{\beta}_x \hat{k}_t - \beta_x^* k_t^* \\
&= \hat{\alpha}_x + \hat{\beta}_x \hat{k}_t - \frac{\hat{\beta}_x}{\sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x} (\hat{k}_t - \bar{k}) \sum_{x=x_{min}}^{x_{max}} \hat{\beta}_x \\
&= \hat{\alpha}_x + \hat{\beta}_x \bar{k}. \tag{3.24}
\end{aligned}$$

A estimação dos parâmetros α_x^* , β_x^* e k_t^* assegura as restrições 3.6 e 3.7 e fornecem a mesma estimativa do número de óbitos $\hat{d}_{x,t}$ dado que $\hat{\alpha}_x + \hat{\beta}_x \hat{k}_t = \alpha_x^* + \beta_x^* k_t^*$.

3.4 Modelação do Nível Geral de Mortalidade

A tendência de mortalidade de um país é capturada pelo parâmetro de variação temporal k_t do modelo LC (Russolillo, 2014).

Depois de ajustar o modelo demográfico, é necessário modelar o índice de nível geral da mortalidade, k_t . Para isso, Lee e Carter utilizaram a metodologia de Box-Jenkins (identificação-estimação-diagnóstico) à série temporal dos valores do parâmetro k_t . Lee e Carter sugerem a utilização de um modelo ARIMA(0,1,0), ou seja, um modelo de passeio aleatório com um termo constante

$$k_t = \lambda + k_{t-1} + v_t, \quad v_t \sim N(0, \sigma_v^2) \quad (3.25)$$

onde λ representa a tendência (*drift*), cujo estimador de máxima verosimilhança é dado por $\hat{\lambda} = \frac{(\hat{k}_{tmax} - \hat{k}_{tmin})}{T-1}$.

Não obstante, o modelo ARIMA sugerido pelos autores pode não ser o mais adequado para todos os estudos, dessa forma é explicado no próximo capítulo em que consiste um modelo ARIMA e todos os passos que constituem a metodologia Box-Jenkins.

Para além da metodologia Box-Jenkins, neste trabalho também é utilizado o Método de Holt para modelar o nível geral de mortalidade, k_t , que será definido também no próximo capítulo. De realçar que, à semelhança do Método Lee-Carter, também na Extensão do Método LC a modelação de k_t será realizada pelas duas metodologias referidas anteriormente.

3.5 Limitações do método Lee-Carter e novos desenvolvimentos

Como o método LC foi utilizado para estudos em vários países, ou seja, foi aplicado em realidades distintas daquelas para a qual foi originalmente concebido, as suas qualidades foram evidenciadas mas as suas limitações foram questionadas. Este método apresenta várias vantagens como, por exemplo, o

facto de este incorporar um modelo demográfico, que é simultaneamente rico e parcimonioso, com um modelo de séries temporais, onde é possível obter intervalos de confiança para as previsões. Através deste modelo também é possível obter, de forma indireta, estimativas para um horizonte onde ainda não haja informação disponível. A um outro nível, o método LC permite contornar algumas das principais limitações dos métodos paramétricos de projeção. Em particular, o método LC reduz a complexidade do problema, oferece uma solução simples para a obtenção de intervalos de confiança e contorna o problema da irregularidade dos parâmetros (Bravo, 2007).

Passado alguns anos da publicação do método de Lee-Carter, R. Lee (2000) e R. Lee & Miller (2001) fizeram um balanço do comportamento do método LC original e suscitaram algumas questões relativamente ao comportamento temporal quase linear da mortalidade, à verificação da hipótese base do modelo de que o vetor de coeficientes β_x se mantém fixo no tempo, às condições iniciais de previsão do modelo, ao grau de incerteza incorporado pelo modelo e à adequação dos intervalos de previsão. Em relação ao comportamento da mortalidade, verificou-se que esta nem sempre apresentou uma forma quase linear como aquela representada pela trajetória de k_{t_0} no estudo feito nos EUA no período 1900-1989. Portanto, o comportamento da série k_t não reflete uma propriedade fundamental do comportamento histórico da mortalidade. As condições iniciais da previsão do modelo também foram questionadas uma vez que à medida que o horizonte de previsão aumenta, as restantes fontes de incerteza, especificamente os erros na estimação de β_x e os erros de ajustamento do modelo demográfico, $\epsilon_{x,t}$, tornam-se cada vez menos relevantes, sendo dominados pela incerteza associada a k_t . Quanto aos intervalos de previsão surgiram também algumas dúvidas, visto que estes não refletem a incerteza sobre se a especificação do modelo é correta, nem a incerteza se o futuro será semelhante ao passado. Os intervalos são considerados muito estreitos, depreciando a capacidade do modelo para incorporar a incerteza sobre o futuro. Verificou-se também que as condições iniciais para a previsão não são as mais corretas, dado que, para alguns grupos etários, constata-se que existe uma quebra entre o valor observado para a taxa de mortalidade no último ano do período de análise e o primeiro valor da previsão. Esta questão foi analisada por Bell (1997), sendo que se concluiu que o método LC tinha uma melhor *performance* em termos de previsão se se considerasse α_x como o logaritmo da taxa de mortalidade observada mais recente.

Com o decorrer dos anos foram desenvolvidas inúmeras extensões do método LC, tais como:

- Wilmoth (1993), com o objetivo de utilizar o método LC para a previsão da mortalidade por causa de morte, desenvolveu um processo em que o ajustamento do modelo era realizado apenas numa

etapa. O autor propõe que o ajustamento do modelo 3.1 seja feito pelo método dos mínimos quadrados ponderados, utilizando como ponderadores o número de óbitos observados por idade e ano, usando diretamente a SVD ponderada. Outra proposta feita por este autor foi realizar a estimação do modelo aplicando o método da máxima verosimilhança, especificando um modelo probabilístico cujos parâmetros podem ser estimados por este método;

- L. R. Carter (1996) propôs a utilização de modelos estruturais de séries temporais para a modelação do nível geral de mortalidade, possibilitando que os parâmetros da equação de estimação possam variar no tempo;
- R. Lee & Miller (2001) sugeriram a alteração da reestimação do parâmetro k_t , sendo que este seria reestimado de maneira a que o valor estimado para a esperança de vida à nascença seja igual ao respetivo valor observado num determinado ano;
- L. Carter & Prskawetz (2001) questionam a possibilidade do método LC contemplar alterações temporais nos parâmetros α_x e β_x , isto é, a possibilidade dos padrões etários de mortalidade se alterarem ao longo do tempo;
- J. Wilmoth & Valkonen (2002) desenvolveram uma extensão do método LC com a finalidade de analisar os diferenciais de mortalidade entre idades e tempo, onde reúnem um conjunto de variáveis explicativas.

Capítulo 4

Séries Temporais

Neste capítulo são apresentadas algumas noções básicas de séries temporais e processos estocásticos, necessárias para a compreensão deste trabalho.

4.1 Conceito de Série Temporal

Uma série temporal é uma sucessão de observações ordenadas no tempo. As séries temporais podem ser contínuas ($Y_t, t \in \mathbb{R}$) ou discretas ($Y_t, t = 1, 2, \dots, n$) e podem ser também classificadas em univariadas, quando são constituídas por observações de uma só variável, ou em multivariadas, quando se observam mais do que uma variável em cada instante. A análise de séries temporais tem como objetivo descobrir o modelo que melhor descreve o comportamento dos dados, isto é, estabelecer uma equação em que a variabilidade dos dados é explicada a partir:

- de funções determinísticas do tempo;
- do seu próprio passado já observado;
- e, ainda, de outras variáveis que podem apenas representar erros de medição, outras quantidade que influenciam o fenómeno ou simplesmente fatores que desconhecemos e/ou que não podemos controlar.

Segundo alguns autores, uma série temporal pode apresentar quatro componentes de variação: a tendência (T), a componente sazonal (S), a componente cíclica (C) e a componente irregular/residual (E). Estas são descritas como:

- A tendência (T) identifica a inclinação, positiva ou negativa, que certas séries apresentam ao longo do tempo. Esta pode ser consequência do facto dos valores observados dependerem de uma componente determinística que é função monótona do tempo;
- A componente sazonal (S) pode corresponder a um aumento ou decréscimo que ocorre regularmente em períodos especiais do ano, originando oscilações que se repetem;
- A componente cíclica (C) corresponde a um padrão de flutuação que não apresenta uma periodicidade definida, isto é, a sua duração não é fixa e, portanto, o seu comprimento varia frequentemente de ciclo para ciclo, nem apresenta uma causa atribuída a fatores "sazonais";
- A componente irregular/residual (E) é a componente que contém qualquer variação não explicada pelas componentes anteriores e representa o ruído aleatório.

Considere-se que T_t corresponde à tendência no tempo t , S_t à componente sazonal no tempo t , E_t à componente irregular no tempo t e que a componente cíclica (C) encontra-se incorporada na tendência. Alguns modelos habitualmente adotados são:

- **Modelo aditivo**

$$Y_t = T_t + S_t + E_t; \quad (4.1)$$

- **Modelo Multiplicativo**

$$Y_t = T_t \times S_t \times E_t. \quad (4.2)$$

Um modelo aditivo é apropriado quando a magnitude das oscilações sazonais não varia com o nível da série. Contudo, se estas aumentam ou diminuem proporcionalmente com a tendência da série, então o modelo multiplicativo é o mais adequado (Makridakis et al., 1998).

Por vezes é necessário realizar uma transformação dos dados, sendo que a mais utilizada é a logarítmica, de forma a converter um modelo multiplicativo num modelo aditivo, isto é, aplicando a transformação logarítmica em 4.2 obtém-se

$$\log(Y_t) = \log(T_t) + \log(S_t) + \log(E_t). \quad (4.3)$$

Desta forma obtém-se um modelo multiplicativo através do ajustamento de um modelo aditivo ao logaritmo dos dados, porém esta transformação não deve ser aplicada a séries temporais com valores negativos ou nulos.

Os modelos de decomposição aditivo e multiplicativo não são os únicos modelos para decompor uma série temporal, a junção dos dois pode originar outros modelos que incluem relações tanto aditivas como multiplicativas. Por exemplo, o resultado da combinação entre os dois modelos pode originar um modelo multiplicativo com erros aditivos, ou seja,

$$Y_t = T_t \times S_t + E_t. \quad (4.4)$$

A decomposição de séries temporais pode ser feita de forma automática no ambiente R , através da função *decompose* do *package stats*. Esta permite avaliar as diversas componentes de uma série temporal separadamente, ajudando a identificar o comportamento individual das mesmas.

4.2 Processos Estocásticos

Dado um processo estocástico $\{Y(t), t \in \mathcal{T}\}$, uma série temporal é um conjunto de observações do processo estocástico em instantes t_1, t_2, \dots, t_n . Generalizando, considera-se t inteiro (i.e., $t = 0, \pm 1, \pm 2, \dots$) e as observações são feitas em intervalos de tempo regulares, isto é, com a mesma amplitude (Alpuim, 1998).

Definição 1. *Um processo estocástico é qualquer família ou coleção de variáveis aleatórias $Y(t), t \in \mathcal{T}$, em que $Y(t)$ é uma variável aleatória (ou conjunto de variáveis aleatórias) com contradomínio S , denominado por espaço de estados. \mathcal{T} é um conjunto de índices ordenados representando o tempo (Alpuim & Ribeiro, 2003).*

Ao conjunto de índices \mathcal{T} dá-se o nome de *espaço de parâmetros* e ao contradomínio das variáveis aleatórias $Y(t)$ dá-se o nome de *espaço de estados*, que se representa por S . Quanto à natureza de \mathcal{T} , se $\mathcal{T} = \mathbb{Z}$ ou $\mathcal{T} = \mathbb{N}$ diz-se que o processo é de tempo discreto e se $\mathcal{T} = \mathbb{R}$ ou, mais comumente, $\mathcal{T} = \mathbb{R}^+$ diz-se que o processo é de tempo contínuo.

Existem duas formas de caracterizar um processo estocástico. A primeira é especificar a distribuição de probabilidade conjunta das suas n variáveis aleatórias $(Y(t_1), \dots, Y(t_n))$ para todos os inteiros n e pontos t_1, \dots, t_n . Porém esta forma de caracterizar é de extrema dificuldade e na prática não é viável. Uma forma simples de descrever um processo estocástico é através dos momentos do processo, em particular o primeiro (valor médio) e segundo momento (função de autocovariância), designados por

$$- \mu(t) = E[Y(t)], \text{ para } t = 0, \pm 1, \pm 2, \dots,$$

$$-\gamma(t_1, t_2) = E[(Y(t_1) - \mu(t_1))(Y(t_2) - \mu(t_2))], t = 0, \pm 1, \pm 2, \dots$$

A variância $\sigma^2(t) = Var[Y(t)]$ para $t = 0, \pm 1, \pm 2, \dots$, é um caso particular da função de autocovariância (quando $t_1 = t_2$).

A série de valores observados, que compõem a série temporal, é considerada apenas uma realização (ou trajetória) de um processo estocástico, de entre todas as possíveis, ou seja, o processo estocástico é um só, porém as séries que este pode gerar são infinitas.

Os processos estocásticos podem ser estacionários ou não estacionários. De seguida, são apresentados os dois tipos de estacionariedade (forte e fraca), alguns procedimentos que permitem transformar processos não estacionários em estacionários e outras ferramentas necessárias para realizar a modelação de séries temporais, como as funções de autocorrelação (FAC), funções de autocorrelação parcial (FACP) e o processo de ruído branco.

4.2.1 Processos Estocásticos Estacionários

Nas Definições 2 e 3 estão apresentadas, segundo Menezes (2020), os conceitos de processo estocástico estacionário (ou fortemente estacionário) e processo estocástico de 2ª ordem (ou fracamente estacionário).

Definição 2. Um processo estocástico $Y(t), t \in \mathcal{T}$ diz-se estritamente estacionário se a distribuição conjunta de $(Y(t_1), \dots, Y(t_n))$ é igual à distribuição conjunta de $(Y(t_1 + \delta), \dots, Y(t_n + \delta))$ qualquer que seja o n -úpla (t_1, \dots, t_n) e para qualquer δ , i.e.,

$$F_{(Y(t_1), \dots, Y(t_n))}(y_1, \dots, y_n) = F_{(Y(t_1 + \delta), \dots, Y(t_n + \delta))}(y_1, \dots, y_n) \quad (4.5)$$

em todos os pontos (y_1, \dots, y_n) .

Definição 3. Um processo estocástico $\{Y(t), t \in \mathcal{T}\}$ diz-se estacionário de 2.ª ordem (ou fracamente estacionário) se todos os momentos até à 2.ª ordem de $(Y(t_1), \dots, Y(t_n))$ existem e são iguais aos momentos correspondentes até à 2.ª ordem de $(Y(t_1 + \delta), \dots, Y(t_n + \delta))$. Logo, num processo fracamente estacionário:

1. o valor médio não depende de t , i.e., $\mu(t) = \mu$;
2. a variância não depende de t , i.e., $\sigma^2(t) = \sigma^2$;

3. a covariância de $Y(t_1)$ e $Y(t_2)$ depende apenas do desfasamento $t_2 - t_1$, i.e.,
 $Cov[Y(t_1), Y(t_2)] = \gamma(|t_2 - t_1|)$.

Nos próximos conceitos consideram-se apenas processos estacionários de 2ª ordem, designados por processos estacionários.

4.2.2 Funções de Autocovariância, Autocorrelação e Autocorrelação Parcial

Definição 4. Para um processo estacionário, define-se a função de autocovariância

$$\gamma_s = Cov[Y_t, Y_{t+s}] = E[(Y_t - \mu)(Y_{t+s} - \mu)], \quad (4.6)$$

que mede a intensidade com que covariam (se acompanham) pares de valores do processo separados por um intervalo (lag) de amplitude s .

A função de autocovariância γ_s é definida para $s \in \mathbb{R}$ se o processo é de tempo contínuo e para $s \in \mathbb{Z}$ se for de tempo discreto, i.e., $s = 0, \pm 1, \pm 2, \dots$. Além disso, esta função satisfaz as seguintes propriedades:

1. $\gamma_0 = Cov[Y_t, Y_t] = Var[Y_t] = \sigma^2$;
2. $\gamma_s = \gamma_{-s}$, i.e., a função é par e dispensa a representação gráfica para $s < 0$;
3. $|\gamma_s| \leq \gamma_0$, como consequência da desigualdade de Cauchy-Schwarz dada por $|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$;
4. A função γ_s é semidefinida positiva, i.e., para qualquer conjunto de números reais $\alpha_1, \dots, \alpha_n$ e instantes de tempo t_1, \dots, t_n ,

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \gamma(|t_i - t_j|) \geq 0. \quad (4.7)$$

Definição 5. A função de autocorrelação (FAC), para um processo estacionário, é definida por

$$\rho_s = Corr[Y_t, Y_{t+s}] = \frac{Cov[Y_t, Y_{t+s}]}{\sqrt{Var[Y_t]Var[Y_{t+s}]}} = \frac{Cov[Y_t, Y_{t+s}]}{Var[Y_t]} = \frac{\gamma_s}{\gamma_0}, \quad (4.8)$$

que mede a correlação entre pares de valores do processo separados por um intervalo de amplitude s .

Tal como no caso da função de autocovariância, a função de autocorrelação ρ_s pode estar definida para $s \in \mathbb{R}$ ou para $s \in \mathbb{Z}$, consoante o processo for de tempo contínuo ou discreto, respetivamente. Além disso, esta função satisfaz as seguintes propriedades:

1. $\rho_0 = Corr[Y_t, Y_t] = 1$;
2. $\rho_s = \rho_{-s}$, i.e., a função é par e dispensa a representação gráfica para $k < 0$;
3. $|\rho_s| \leq 1$, como consequência da desigualdade de Cauchy-Schwarz;
4. A função ρ_s é semidefinida positiva, i.e., para qualquer conjunto de números reais $\alpha_1, \dots, \alpha_n$ e instantes de tempo t_1, \dots, t_n ,

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \rho(|t_i - t_j|) \geq 0. \quad (4.9)$$

O correlograma teórico é a representação gráfica da FAC. A análise deste ajuda na identificação de várias características de uma série temporal e é um auxiliar na escolha do modelo que lhe é mais adequado. O aumento de s traduz-se no decréscimo de ϕ_s e, conseqüentemente, de γ_s , isto é, a memória desvanece ao longo do tempo. Já quando s tende para $+\infty$, ϕ_s e γ_s tendem ambos para zero.

Para além de estudar a correlação de uma forma global, por vezes também interessa estudar a correlação parcial entre Y_t e Y_{t+s} , quando se fixam as variáveis intermédias $Y_{t+1}, Y_{t+2}, \dots, Y_{t+s-1}$.

Definição 6. O conjunto de autocorrelações parciais de desfaseamento (lag) s é dado por $\{\phi_{ss} : s = 1, 2, \dots\}$ onde

$$\phi_{ss} = Corr[X_t, X_{t+s} | X_{t+1}, X_{t+2}, \dots, X_{t+s-1}] = \frac{|P_s^*|}{|P_s|} \quad (4.10)$$

e P_s^* é a matriz $s \times s$ de autocorrelações onde a última coluna é substituída por $[\rho_1 \ \rho_2 \ \dots \ \rho_s]^T$. A matriz P_s é dada por

$$P_s = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{s-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{s-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{s-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{s-1} & \rho_{s-2} & \cdots & \rho_1 & 1 \end{bmatrix}. \quad (4.11)$$

Esta função satisfaz as seguintes propriedades:

1. $\phi_{11} = \rho_1$;

$$2. \phi_{22} = \frac{\phi_2 - \phi_1^2}{1 - \phi_1^2};$$

$$3. \phi_{33} = \frac{\rho_3(1 - \rho_1^2) + \rho_1(\rho_1^2 + \rho_2^2 - 2\rho_2)}{(1 - \rho_2)(1 + \rho_2 - 2\rho_1^2)}.$$

4.2.3 Ruído Branco

Definição 7. Um processo estocástico $\{\epsilon_t, t \in \mathbb{Z}\}$ diz-se um processo puramente aleatório ou processo de ruído branco quando é formado por uma sucessão de variáveis aleatórias não correlacionadas e identicamente distribuídas, de média e variância constantes, i.e., um processo estocástico diz-se um ruído branco se e só se satisfaz as seguintes condições:

1. $E[\epsilon_t] = \mu_\epsilon$ (usualmente $\mu_\epsilon = 0$);
2. $Var[\epsilon_t] = \sigma_\epsilon^2$;
3. $Cov(\epsilon_t, \epsilon_{t+s}) = \gamma_s = 0, s = \pm 1, \pm 2, \dots$

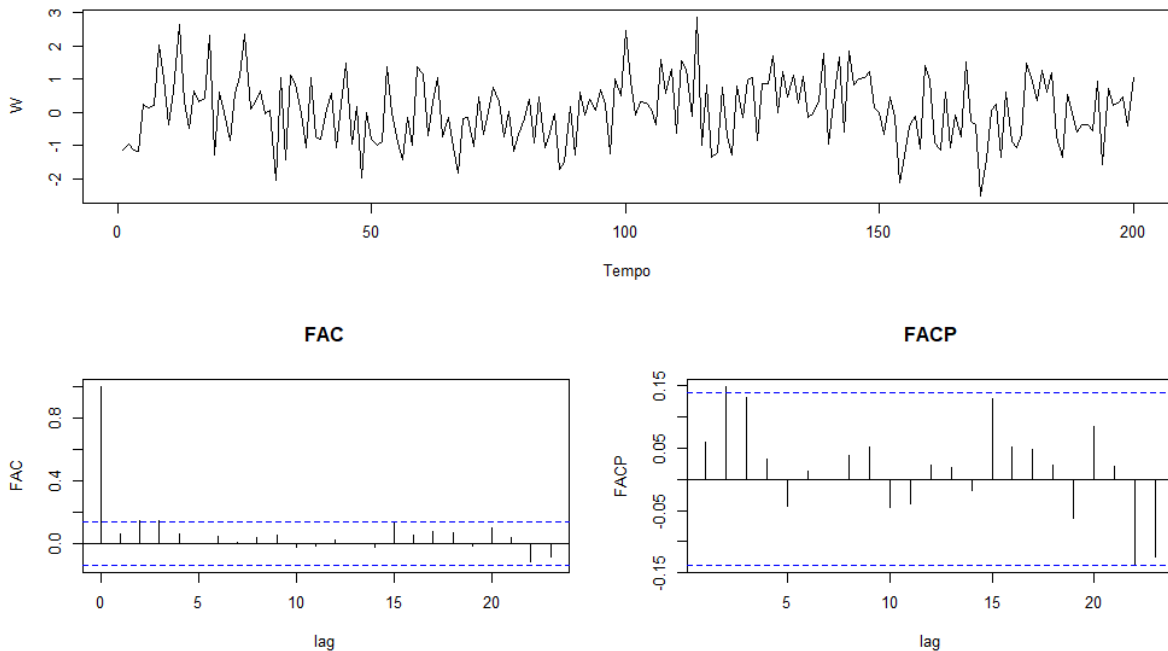


Figura 4.1: Simulação de um ruído branco e respetivas FAC e FACP empíricas.

Como a média e a função de autocovariância não dependem do tempo, o processo é estacionário de segunda ordem. Se para além disso as variáveis aleatórias seguirem uma distribuição Normal (i.e.,

$\epsilon_t \sim N(\mu_\epsilon, \sigma_\epsilon^2)$, então o processo é designado de ruído branco Gaussiano. Um bom modelo de previsão deverá produzir erros de previsão com um comportamento análogo ao de um ruído branco, dada a imprevisibilidade inerente ao ruído branco. A título de exemplo, na Figura 4.1 é apresentada uma simulação de um processo de ruído branco e as respetivas FAC e FACP empíricas.

4.2.4 Estimação dos parâmetros caracterizadores dos processos estacionários

Na modelação de uma série temporal é extremamente importante a estimação de parâmetros que caracterizem o processo estacionário subjacente. Os processos estacionários são caracterizados pela média μ , a variância σ^2 e as autocorrelações ρ_s . Considerando um conjunto de n observações de um processo estacionário Y_t durante um certo período de tempo, ou seja, Y_1, Y_2, \dots, Y_n , podem utilizar-se os estimadores clássicos dos parâmetros

- para a média μ usar $\bar{Y} = \frac{1}{n} \sum_{t=1}^n Y_t$;
- para a autocovariância γ_s usar $\hat{\gamma}_s = \frac{1}{n} \sum_{t=1}^{n-s} (Y_t - \bar{Y})(Y_{t+s} - \bar{Y})$;
- para a autocorrelação ρ_s usar $\hat{\rho}_s = \frac{\hat{\gamma}_s}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-s} (Y_t - \bar{Y})(Y_{t+s} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$.

4.3 Processos Estocásticos Não Estacionários

Numa série estacionária, os valores futuros serão similares aos do passado e, por isso, a estacionariedade é um importante pressuposto quando se pretende obter previsões com base em observações passadas. A maioria das séries temporais, na prática, são não estacionárias. Um processo pode ser não estacionário por a média e/ou a variância serem funções do tempo e não constantes. Uma série estacionária em média pode não ser estacionária em variância. Para contornar este problema pode recorrer-se a transformações que estabilizam a média e/ou variância, de forma a converter uma série temporal não estacionária numa série estacionária. No caso de se estar perante uma série não estacionária em média e em variância, deve se proceder, primeiramente, à estabilização da variância e, posteriormente, à estabilização da média.

Existem vários procedimentos que têm como objetivo a remoção da tendência e da sazonalidade de uma série temporal, permitindo que se atinja a estacionariedade. Estes métodos seguem os modelos de decomposição abordados anteriormente e consistem na estimação das componentes Tendência (T_t) e sazonalidade (S_t) através de funções determinísticas (ou outras abordagens), de modo que a série, após a "remoção" dessas componentes, passe a ser bem modelada por um processo estacionário (Alpuim, 1998). Porém, em muitas séries temporais é possível realizar outro tipo de transformações que permitem transformar séries não estacionárias em estacionárias.

4.3.1 Transformações para a estacionariedade

Uma das formas que permite a estabilização da média consiste no uso de processos de diferenciação, que resultam da aplicação do operador diferença ∇ , definido como $\nabla Y_t = Y_t - Y_{t-1}$, à série temporal não estacionária. Assim, se uma série, Y_t , for não estacionária, pode pensar-se em transformá-la numa série estacionária, aplicando uma diferenciação de primeira ordem, i.e.,

$$\nabla Y_t = Y_t - Y_{t-1}, \quad t = 2, 3, \dots, n. \quad (4.12)$$

Se a diferenciação de 1.^a ordem não for suficiente para obter uma série estacionária, podem obter-se as diferenças de 2.^a ordem, que correspondem às diferenças das primeiras diferenças da série original, i.e.,

$$\nabla^2 Y_t = \nabla(\nabla Y_t) = \nabla(Y_t - Y_{t-1}) = Y_t - 2Y_{t-1} + Y_{t-2}, \quad t = 3, 4, \dots, n. \quad (4.13)$$

O operador de diferenciação de ordem d , para qualquer inteiro $d \geq 1$, consiste em diferenciar a série d vezes, ou seja,

$$\nabla^d Y_t = \nabla(\nabla^{d-1} Y_t), \quad t = d + 1, \dots, n. \quad (4.14)$$

De notar que a diferenciação de uma série já estacionária é indesejável e deve evitar-se. O objetivo passa por determinar a série estacionária obtida pela menor diferenciação, uma vez que a variância aumenta com a ordem de diferenciação. De uma forma geral, se a série transformada $\nabla^{d_0} Y_t$ é estacionária, então, para qualquer $d > d_0$, a série $\nabla^d Y_t$ é também estacionária, mas tem maior variância. Conclui-se, portanto, que se deve evitar a sobrediferenciação para não introduzir variação indesejada na série transformada. Caso este procedimento seja adequado, as diferenciações de 1.^a e 2.^a ordens devem ser suficientes para obter uma série estacionária.

Se uma série temporal é diferenciada uma vez e a série diferenciada é estacionária, diz-se que a série original é integrável de ordem 1 e representa-se por $I(1)$. Usualmente, se a série for diferenciada d vezes, esta é integrável de ordem d ou $I(d)$. Por convenção, se $d = 0$, o processo $I(0)$ é um processo estacionário.

Para a estabilização da variância de uma série não estacionária, pode ser utilizado um método de transformação paramétrica, mais conhecido como transformação de Box-Cox, que se baseia na seguinte expressão

$$Z_t = T(Y_t) = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log Y_t, & \lambda = 0 \end{cases}, \quad (4.15)$$

onde os valores de λ estão no intervalo $[-1; 1]$. As transformações mais recorrentes encontram-se sumariadas na Tabela 4.1.

Tabela 4.1: Transformações usuais de Box-Cox.

λ	Transformação
-1	$Z_t = \frac{1}{Y_t}$
-0,5	$Z_t = \frac{1}{\sqrt{Y_t}}$
0	$Z_t = \log Y_t$
0,5	$Z_t = \sqrt{Y_t}$
1	$Z_t = Y_t$

Porém, note-se que algumas destas transformações, nomeadamente a do logaritmo, apenas estão definidas para séries temporais de valores positivos. Para contornar esta situação, pode-se encontrar uma constante c tal que $Y_t + c > 0$ e, posteriormente, aplicar as transformações.

Como foi referido anteriormente, a transformação logarítmica permite converter o efeito sazonal multiplicativo em aditivo, sendo por isso uma das transformações Box-Cox mais utilizadas. Esta transformação permite também estabilizar as diferenças entre os valores da série temporal, isto é, a sua variância, e pode ser utilizada para atenuar o efeito de possíveis *outliers*. De realçar que, sempre que se aplica uma transformação aos dados, todas as previsões geradas pelo modelo selecionado estarão nas unidades transformadas. Por essa razão, depois de ajustado o modelo e estimados os parâmetros, devem ser revertidas as transformações, de modo a obter previsões nas unidades originais.

Passeio aleatório

Considere-se o seguinte modelo de tendência estocástica,

$$Y_t = Y_{t-1} + \epsilon_t, \quad (4.16)$$

onde ϵ_t é um ruído branco. Este modelo é vulgarmente conhecido por modelo de passeio aleatório (*random walk*), uma vez que descreve a série no momento t com base no seu valor passado (no momento $t - 1$) mais um choque aleatório.

Em geral, os “passeios aleatórios” caracterizam-se por movimentos de tendência crescente ou decrescente em períodos longos, seguidos de mudanças bruscas de sentido imprevisível (Caiado, 2022). No entanto, uma vez que as aparentes tendências são apenas o resultado da existência de um elevado número de valores positivos em relação ao número de valores negativos (no caso de tendências crescentes) ou vice-versa (no caso de tendências decrescentes), o valor médio de um processo aleatório é constante. Efetivamente, se Y_0 for conhecido, o modelo de passeio aleatório pode ser reescrito como

$$Y_t = Y_0 + \sum_{i=1}^t \epsilon_i, \quad (4.17)$$

e, calculando o valor esperado de Y_t , conclui-se que este é igual a Y_0 e, portanto, constante. Apesar disso, o modelo de passeio aleatório é um processo não estacionário, uma vez que a sua variância depende de t (Enders, 2015). No entanto, aplicando uma transformação, nomeadamente uma diferenciação de 1.^a ordem, $\nabla Y_t = \epsilon_t$, é possível obter um ruído branco, que é um processo estacionário.

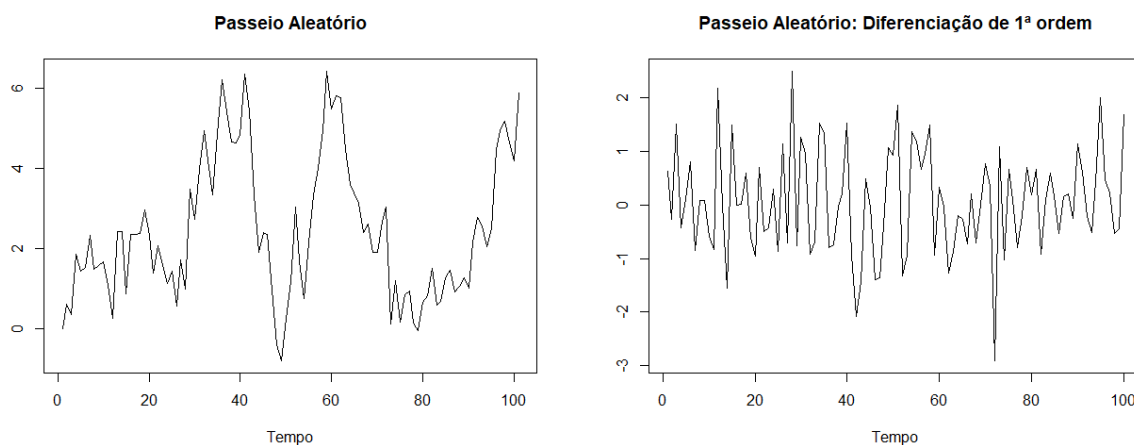


Figura 4.2: Simulação de um modelo de passeio aleatório e as suas diferenças de 1.^a ordem.

Na Figura 4.2 pode observar-se uma série simulada com 100 observações de um modelo de passeio aleatório, com ϵ_t um ruído branco Gaussiano de média nula e variância unitária, bem como o resultado

da diferenciação de 1.^a ordem da mesma. O correlograma de um passeio aleatório apresenta correlações positivas que decaem muito lentamente para zero.

Passeio aleatório com drift

O modelo de passeio aleatório com *drift* é uma extensão do modelo anterior, quando se adiciona um termo constante a_0 , i.e.,

$$Y_t = a_0 + Y_{t-1} + \epsilon_t, \quad (4.18)$$

onde ϵ_t é um ruído branco. Neste caso, considera-se que a tendência é parcialmente determinística e parcialmente estocástica. Por essa razão, o valor médio deste processo depende de t , não sendo constante. De facto, se Y_0 for conhecido, o modelo de passeio aleatório com *drift* pode ser reescrito como

$$Y_t = Y_0 + a_0 t + \sum_{i=1}^t \epsilon_i, \quad (4.19)$$

e, calculando o valor esperado de Y_t , conclui-se que este é igual a $Y_0 + a_0 t$ e, por isso, dependente de t . Desta forma, também o modelo de passeio aleatório com *drift* é não estacionário, podendo ser transformado em estacionário se se aplicar uma diferenciação de 1.^a ordem.

Na Figura 4.3 pode observar-se uma série simulada com 100 observações de um modelo de passeio aleatório com *drift*, onde ϵ_t é um ruído branco Gaussiano de média nula e variância unitária e $a_0 = 0,5$, bem como o resultado da diferenciação de 1.^a ordem da mesma.

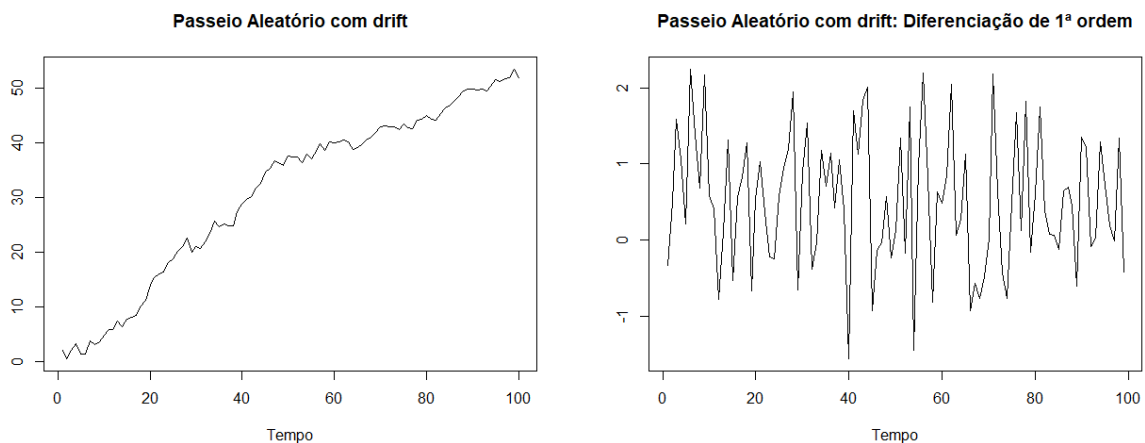


Figura 4.3: Simulação de um modelo de passeio aleatório com *drift* e as suas diferenças de 1.^a ordem.

Usualmente, neste tipo de processos, a tendência determinística é dominante. No entanto, o aumento

da variância de ϵ_t ou a diminuição do valor absoluto de a_0 podem mascarar a existência de uma tendência determinística (Enders, 2015).

4.3.2 Análise de estacionariedade

Para analisar a estacionariedade de uma série temporal, de uma forma, elementar, pode-se representar graficamente os dados ao longo do tempo. Contudo, esta análise é subjetiva e, apesar de útil, deve ser confirmada através de testes estatísticos formais. Existem diversos testes para realizar este estudo, estes baseiam-se, na sua maioria, em encontrar uma raiz unitária. De acordo com a literatura, deve-se utilizar mais do que um teste de raiz unitária, de forma a avaliar a estacionariedade da série. Alguns dos testes mais utilizados são os testes de Dickey-Fuller (DF), Dickey-Fuller Aumentado (*Augmented Dickey Fuller*, ADF), Phillips-Perron (PP) e Kwiatkowski-Phillips-Schmidt-Shin (KPSS). No caso dos testes DF, ADF e PP, a hipótese a testar é a presença de uma raiz unitária, ou seja, a não estacionariedade, já no caso do teste KPSS, a hipótese nula é a estacionariedade.

Teste de Dickey-Fuller

Considere-se um processo definido por

$$Y_t = \phi Y_{t-1} + \epsilon_t, \quad -1 \leq \phi \leq 1, \quad (4.20)$$

onde ϵ_t é um ruído branco. Este processo é estacionário se $|\phi| < 1$ (trata-se de um processo autorregressivo de ordem 1, como se verá na Secção 5.2). Porém, quando $\phi = 1$ este processo é um passeio aleatório, o que, como foi visto anteriormente, equivale a fazer uma diferenciação de 1.ª ordem. Assim, se $\phi = 1$, pode dizer-se que a série é não estacionária.

O processo descrito pela equação (4.20) pode ser escrito na forma das diferenças, ou seja,

$$Y_t = \phi Y_{t-1} + \epsilon_t \Leftrightarrow Y_t - Y_{t-1} = \phi Y_{t-1} - Y_{t-1} + \epsilon_t \quad (4.21)$$

$$\Leftrightarrow \nabla Y_t = (\phi - 1)Y_{t-1} + \epsilon_t \quad (4.22)$$

$$\Leftrightarrow \nabla Y_t = \delta Y_{t-1} + \epsilon_t, \quad (4.23)$$

onde $\delta = \phi - 1$ e ϵ_t é um processo estacionário. Assim, as hipóteses para o teste DF em termos de δ são dadas por

$$H_0 : \delta = 0 \text{ vs } H_1 : -2 < \delta < 0.$$

De notar que, se $\delta = 0$ então $Y_t \sim I(1)$. Isto é, a série das primeiras diferenças é estacionária e a série original é um passeio aleatório. Se $-2 < \delta < 0$, então o processo Y_t é estacionário. Com base nesta modificação, Dickey & Fuller (1979) propõem o teste

$$H_0 : \delta = 0 \text{ vs } H_1 : \delta < 0.$$

Se a hipótese H_0 não é rejeitada então o processo Y_t é não estacionário, i.e., $Y_t \sim I(1)$. No caso de $\delta < 0$ então $\phi - 1 < 0 \Leftrightarrow \phi < 1$ e, tendo em conta as condições de estacionariedade para este processo, a série Y_t é estacionária.

Além da equação (4.23), Dickey e Fuller consideram também duas equações alternativas:

$$\nabla Y_t = a_0 + \delta Y_{t-1} + \epsilon_t, \quad (4.24)$$

$$\nabla Y_t = a_0 + a_1 t + \delta Y_{t-1} + \epsilon_t. \quad (4.25)$$

A diferença entre as três equações reside, essencialmente, na presença/ausência de uma constante a_0 e/ou de um termo determinístico $a_1 t$. Assim, a rejeição da hipótese nula, $H_0 : \delta = 0$, considerando a equação (4.24), implica que o processo Y_t é estacionário de média não nula. No caso da equação (4.25), a indicação é de que $Y_t \sim I(0)$ com uma tendência determinística.

Ao aplicar o teste DF deve verificar-se se a sequência ϵ_t é não correlacionada, uma vez que, quando este pressuposto não se verifica, estes testes tendem a exibir um fraco comportamento. Caso isso aconteça, o teste DF não pode ser aplicado e deve optar-se por uma versão do teste para ordens superiores, ou seja, o teste ADF (Cordeiro, 2011).

Teste de Dickey-Fuller Aumentado

Muitas séries apresentam uma estrutura mais complexa do que a captada pelo modelo (4.20). Para lidar com esses casos, surge uma variante do teste DF capaz de integrar modelos mais complexos. Esta inovação deve-se a Said & Dickey (1984) e, desde então, este tem sido um dos testes mais usados no estudo da estacionariedade de séries temporais.

Considere-se um processo mais complexo que o apresentado em (4.20) definido por

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t. \quad (4.26)$$

Assim, repetindo o processo da (4.21) a (4.23), obtém-se

$$\nabla Y_t = \delta Y_{t-1} + \sum_{j=1}^{p-1} \gamma_j \nabla Y_{t-j} + \epsilon_t, \quad (4.27)$$

onde $\delta = \sum_{i=1}^p \phi_i - 1$, $\gamma_j = -\sum_{i=j+1}^p \phi_i$, $\nabla Y_{t-j} = Y_{t-j} - Y_{t-j-1}$ e ϵ_t é um ruído branco. Esta decomposição separa o modelo (4.26) em dois termos: Y_{t-1} e as $p-1$ primeiras diferenças. No caso em que Y_t é $I(1)$ (passeio aleatório), esta separação envolve uma componente $I(1)$ e $p-1$ componentes $I(0)$ (estacionárias). Na terminologia comum, diz-se que o modelo original foi aumentado por $p-1$ componentes de primeiras diferenças, dando origem à designação $ADF(p-1)$.

Para o modelo (4.27), a existência de uma raiz unitária é garantida se $\sum_{i=1}^p \phi_i = 1$, i.e., se $\delta = 0$. Assim, as hipóteses a testar são, também neste caso,

$$H_0 : \delta = 0 \text{ vs } H_1 : \delta < 0$$

e a não rejeição da hipótese nula implica a não estacionariedade de Y_t . Uma vez que este teste é unilateral à esquerda, a hipótese nula é rejeitada a um nível de significância α se a estatística de teste for inferior ou igual ao quantil $(1-\alpha)100\%$ da distribuição correspondente (valor crítico).

Além da equação (4.27) e à semelhança do que acontece no teste DF, são propostas duas outras equações

$$\nabla Y_t = a_0 + \delta Y_{t-1} + \sum_{j=1}^{p-1} \gamma_j \nabla Y_{t-j} + \epsilon_t, \quad (4.28)$$

$$\nabla Y_t = a_0 + a_1 t + \delta Y_{t-1} + \sum_{j=1}^{p-1} \gamma_j \nabla Y_{t-j} + \epsilon_t, \quad (4.29)$$

cuja diferença reside, novamente, na presença/ausência de uma constante a_0 e/ou de um termo determinístico $a_1 t$.

O teste ADF apresenta um problema que é a decisão de qual o número de termos a incluir na equação a ser testada, ou seja, o valor de p . Para isso podem ser utilizadas várias estratégias, tais como:

1. Escolher o valor de p empiricamente. A ideia é incluir um número suficiente de termos de forma que a estrutura dos resíduos não apresente correlação;
2. Minimizar algum critério de informação como, por exemplo, o critério de informação de Akaike (*Akaike Information Criterion*, AIC);
3. Usar a regra proposta por Ng & Perron (1995): primeiro, define-se um limite máximo para p , p_{max} ; de seguida, calcula-se o teste ADF considerando $p = p_{max}$; se o valor absoluto da estatística t para testar a significância da diferença de ordem p for maior do que 1,6, define-se $p = p_{max}$

e prossegue-se com o teste; caso contrário, reduz-se p em uma unidade e repete-se o processo.

Para determinar p_{max} pode utilizar-se (proposta de Schwert (2002))

$$p_{max} = \left[12 \left(\frac{T}{100} \right)^{1/4} \right],$$

onde $[x]$ representa a parte inteira de x e T o número de observações. Esta será a estratégia adotada para definir o valor de p .

Independentemente da estratégia escolhida, o importante é averiguar se os resíduos do modelo apresentam um comportamento análogo ao de um ruído branco antes da aplicação do teste.

Teste de Phillips-Perron

O teste de Phillips-Perron (PP) difere do teste ADF principalmente na estrutura permitida para os erros, ϵ_t . Efetivamente, enquanto que o teste ADF exige que estes sejam independentes, o teste PP permite que os erros sejam correlacionados e, possivelmente, heterocedásticos. Por isso, as equações a testar são iguais às apresentadas para o teste ADF, com a única diferença que ϵ_t deixa de ser um ruído branco. As hipóteses a testar são também semelhantes e, por isso, a rejeição da hipótese nula implica a estacionariedade do processo.

Teste de KPSS

O teste de KPSS, proposto por Kwiatkowski, Phillips, Schmidt e Shin, serve, tal como os anteriores, para avaliar a estacionariedade de um processo. No entanto, neste caso, as hipóteses a testar são diferentes, ou seja,

$$H_0 : \text{O processo é estacionário} \quad vs \quad H_1 : \text{O processo é não estacionário.}$$

Se se considerar o processo Y_t , a equação deste teste decompõe Y_t numa soma de três componentes: uma tendência determinística (T_t), um passeio aleatório (μ_t) e um erro estacionário (u_t), i.e.,

$$Y_t = T_t + \mu_t + u_t, \tag{4.30}$$

$$\mu_t = \mu_{t-1} + \epsilon_t, \tag{4.31}$$

onde ϵ_t é um ruído branco. Este teste é unilateral à direita logo, a hipótese nula é rejeitada a um nível de significância α se a estatística de teste for superior ou igual ao quantil $(1 - \alpha)100\%$ da distribuição correspondente (valor crítico).

Capítulo 5

Modelos de Previsão em Séries Temporais

Um modelo estatístico determina um processo gerador dos dados, que pode ser utilizado para obter toda a distribuição de probabilidade para um momento futuro. Um modelo permite realizar previsões pontuais para um determinado horizonte temporal e também permite o cálculo de previsões intervalares, isto é, intervalos de previsão, para um nível de confiança associado.

Um método de previsão é um procedimento para calcular previsões a partir de valores presentes e passados. Como tal, pode ser simplesmente um algoritmo e não depender de um modelo de probabilidade subjacente ou, alternativamente, surgir da identificação de um modelo específico para os dados fornecidos e da localização de previsões condicionadas a esse modelo (Chatfield, 2000). A escolha do método pode depender de muitas considerações como, por exemplo, do objetivo do cálculo das previsões, do tipo de série temporal e a presença/ausência de componentes como a tendência e/ou a sazonalidade, a dimensão da série temporal (n), o horizonte de previsão (h), o conhecimento e a experiência do analista e, por fim, da disponibilidade dos programas informáticos.

De facto, existem diversos métodos de previsão que podem ser aplicados a uma série temporal, apresentando cada um as suas capacidades e limitações. Nesta dissertação são estudadas duas abordagens distintas: a metodologia clássica de Box-Jenkins e a não paramétrica de Holt.

5.1 Metodologia Box-Jenkins

Box & Jenkins (1970) introduziram os modelos SARIMA (*Seasonal Autoregressive Integrated Moving Average*) de forma a modelar e prever séries temporais estacionárias e não estacionárias. Os modelos SARIMA podem ser divididos em modelos mais simples, isto é, podem ser divididos em: modelo

autorregressivo (*autoregressive*, AR), que considera que o comportamento da série pode ser explicado através do seu passado; modelo de médias móveis (*moving average*, MA), que considera que a série temporal é explicada a partir de uma sucessão de choques aleatórios; e, por último, o modelo autorregressivo e de médias móveis (*autoregressive moving average*, ARMA), que se trata de uma combinação dos dois modelos anteriores. Estes modelos apenas são úteis para séries estacionárias, sendo que no caso de séries não estacionária são utilizados os modelos integrados (ARIMA) ou, para séries que apresentem sazonalidade, os equivalentes sazonais (SARIMA).

A metodologia Box-Jenkins é um processo iterativo constituído por três fases, a identificação, estimação e diagnóstico. Na primeira fase pretende-se identificar o modelo ARIMA mais apropriado através da análise da representação gráfica dos dados e das respetivas FAC e FACP empíricas. Na segunda fase os parâmetros do modelo escolhido são estimados e, por fim, na terceira fase os parâmetros são avaliados quanto à sua significância e é também avaliado o comportamento dos resíduos, mais concretamente se estes apresentam um comportamento análogo ao de um ruído branco. Estas três fases são aplicadas iterativamente até que o modelo final não possa ser melhorado.

Comparativamente com os métodos de alisamento exponencial, a maior vantagem dos modelos ARIMA é que estes têm em consideração uma das principais características dos dados de séries temporais, a dependência temporal, ou seja, a autocorrelação. Desta forma, afirma-se que os modelos ARIMA são apropriados quando se pode assumir que existe algum tipo de relação entre o passado e o futuro, contudo, isso torna-os pouco recomendados para realizar previsões a longo prazo e/ou previsões de séries com mudanças bruscas de comportamento.

5.2 Modelos de Processos Estacionários

Considera-se que uma série estacionária fica completamente definida pelas suas funções média, variância e de autocorrelação. Assim sendo, com a metodologia Box-Jenkins, pretende-se identificar um modelo com base no comportamento da função de autocorrelação empírica.

Os processos ARMA são considerados como um grupo bastante variado e de grande fiabilidade no que respeita à modelação de inúmeras séries temporais estacionárias. Contudo, para a modelação de séries que apresentem oscilações bruscas ao longo do tempo, este tipo de processos torna-se insuficiente.

5.2.1 Processo Autorregressivo de ordem p , $AR(p)$

Os processos autorregressivos baseiam-se no pressuposto de que a observação da variável no instante t se relaciona, de forma linear, com as observações nos instantes anteriores. Assim, o processo Y_t diz-se um processo autorregressivo de ordem p , $AR(p)$, quando satisfaz a equação

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_{t-p} Y_{t-p} + \epsilon_t, \quad (5.1)$$

onde ϵ_t é um ruído branco de média nula, independente de Y_{t-s} para todo o $s \geq 1$. De facto, Y_t pode ser vista como uma variável dependente que é explicada através de uma regressão linear múltipla, em que as observações em p instantes anteriores funcionam como variáveis explicativas e ϕ_i são os coeficientes de cada Y_{t-i} .

De forma alternativa, a representação de um processo $AR(p)$ pode ser feita a partir do operador atraso B^s , que se define como sendo $B^s Y_t = Y_{t-s}$. Com efeito, a equação (5.1) pode ser reescrita como

$$\Phi_p(B) Y_t = \epsilon_t, \quad (5.2)$$

onde $\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o polinómio autorregressivo de ordem p . Tendo em consideração as p raízes (reais ou complexas), $G_1^{-1}, G_2^{-1}, \dots, G_p^{-1}$, da equação característica $\Phi_p(B) = 0$, torna-se possível fatorizar o polinómio autorregressivo do seguinte modo

$$\Phi_p(B) = \prod_{i=1}^p (1 - G_i B). \quad (5.3)$$

Para que o processo seja estacionário é condição necessária e suficiente que as raízes da equação característica sejam todas de módulo maior do que a unidade, ou seja, que $|G_i| < 1$, para $i = 1, 2, \dots, p$. Qualquer processo autorregressivo que seja estacionário é também invertível, o que, em termos práticos, significa que a dependência do passado se vai atenuando à medida que o passado se torna mais remoto.

Portanto, se o processo Y_t é um processo $AR(p)$, então a sua função de autocorrelação parcial, ϕ_{ss} , é igual a zero para todo o $s > p$. Assim, a FACP de um processo $AR(p)$ apresenta, graficamente, uma queda brusca para zero a partir do $\text{lag } p + 1$, enquanto que a respetiva FAC tem um decaimento exponencial ou sinusoidal amortecido para zero. Na Figura 5.1 encontra-se representado um processo autorregressivo de ordem 2, $AR(2)$, e as respetivas FAC e FACP empíricas.

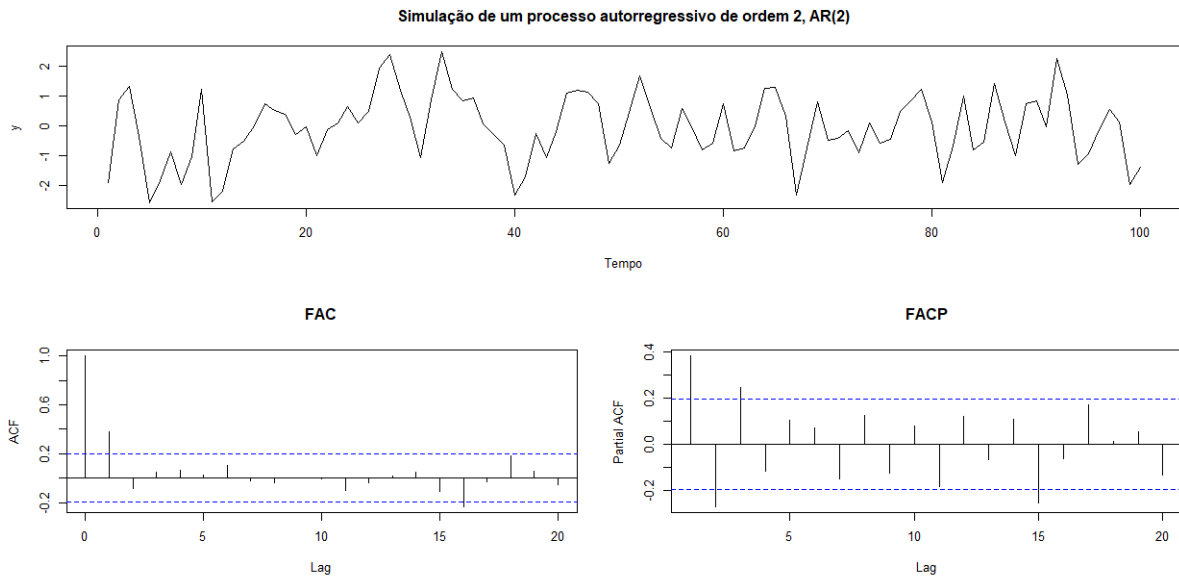


Figura 5.1: Simulação de um processo autorregressivo e as respetivas FAC e FACP empíricas.

5.2.2 Processo de Médias Móveis (MA)

Diz-se que o processo Y_t é um processo de médias móveis de ordem q , $MA(q)$, quando este assume a expressão

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} \quad (5.4)$$

ou

$$Y_t = \Theta_q(B) \epsilon_t, \quad (5.5)$$

onde ϵ_t é um ruído branco de média nula e $\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ é o polinómio de médias móveis de ordem q . Através destes processos, pretende-se exprimir Y_t em termos de um processo mais simples, como é o ruído branco. Desse modo, um processo de médias móveis de ordem q define-se, em cada instante t , como a média ponderada das $q + 1$ observações de um processo de ruído branco. Assim, graças à estacionariedade inerente ao ruído branco, os processos de médias móveis são sempre estacionários. Um processo de médias móveis é invertível se puder ser escrito como um processo autorregressivo estacionário de ordem infinita. Para garantir a invertibilidade do processo, basta que, à semelhança do que acontece no caso da estacionariedade de processos autorregressivos, as raízes da equação característica $\Theta_q(B) = 0$ se encontrem todas fora do círculo unitário, isto é, sejam, em módulo, todas superiores a 1 (Covpertwait & Metcalfe, 2009).

Se o processo Y_t é um processo $MA(q)$, então a sua função de autocorrelação, ρ_s , é igual a zero para

todo o $s > q$, e, por isso, a FAC de um processo $MA(q)$ apresenta, graficamente, uma queda brusca para zero a partir do $lag\ q + 1$. No que respeita à FACP, esta exibe um decaimento exponencial ou sinusoidal amortecido para zero tendo, portanto, a mesma estrutura que a FAC de um processo $AR(q)$. Na Figura 5.2 encontra-se representado um processo de médias móveis de ordem 2, $MA(2)$, e as respetivas FAC e FACP empíricas.

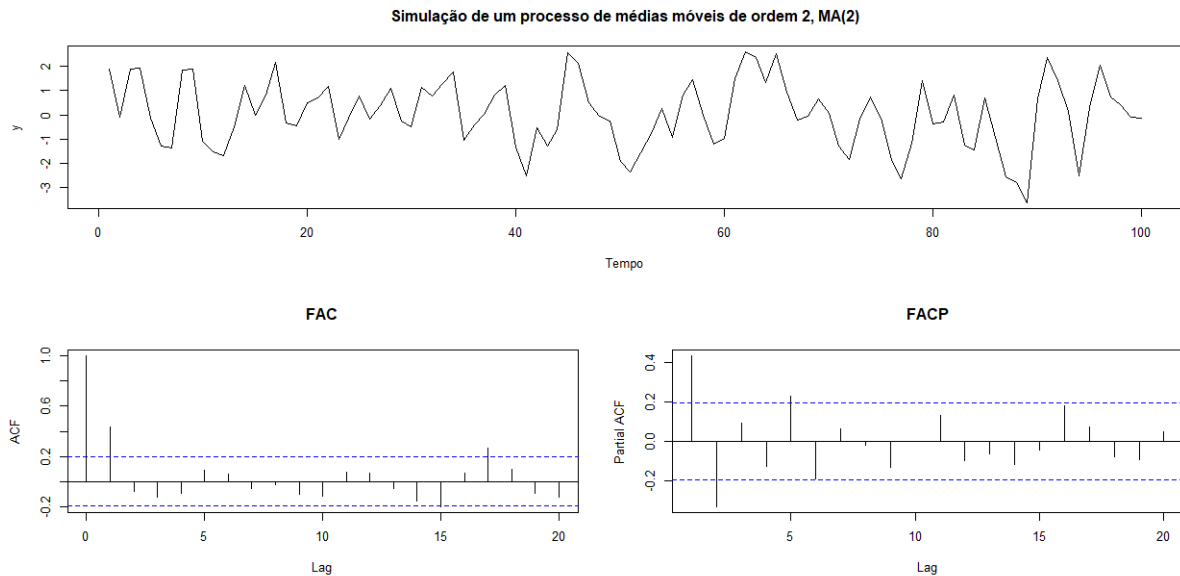


Figura 5.2: Simulação de um processo de médias móveis e as respetivas FAC e FACP empíricas.

5.2.3 Processo Autorregressivo e de Médias Móveis, $ARMA(p, q)$

Os processos estacionários e invertíveis podem ser representados de duas formas, através da forma autorregressiva ou de médias móveis. Porém, pode acontecer que qualquer um destes processos tenha uma representação com um número excessivo de parâmetros, o que pode conduzir a uma perda de eficiência na sua estimação (Caiado, 2022). Se assim se verificar, pode construir-se um modelo mais parcimonioso que inclua tanto termos autorregressivos como de médias móveis. Este modelo corresponde ao modelo $ARMA(p, q)$ que se designa por processo misto autorregressivo e de médias móveis de ordens p e q .

Assim, o processo Y_t diz-se um processo autorregressivo e de médias móveis de ordens p e q , $ARMA(p, q)$, se satisfaz a equação

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (5.6)$$

ou a equação

$$\Phi_p(B)Y_t = \Theta_q(B)\epsilon_t, \quad (5.7)$$

onde ϵ_t é um ruído branco de média nula, independente de Y_{t-s} para todo o $s \geq 1$, $\Phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ e $\Theta_q(B) = 1 + \theta_1 B + \dots + \theta_q B^q$ são os polinómios autorregressivos e de médias móveis de ordens p e q , respetivamente.

A FAC e a FACP de um processo ARMA(p, q) resultam da combinação das respetivas funções dos processos AR(p) e MA(q). Recorde-se que a FAC de um processo MA(q) é insignificante a partir do lag $q + 1$, o mesmo acontecendo para a FACP de um processo AR(p) depois do lag p . De facto, uma vez que o processo ARMA(p, q) é uma combinação dos processos AR(p) e MA(q), a estacionariedade e a invertibilidade do processo ficam garantidas se as raízes das equações características $\Phi_p(B) = 0$ e $\Theta_q(B) = 0$ são, em módulo, maiores do que a unidade. Como é evidente, estes processos generalizam os processos anteriormente mencionados e, por exemplo, um processo ARMA($p, 0$) é equivalente a um processo AR(p), o mesmo acontecendo com um ARMA($0, q$) relativamente a um MA(q).

Na Figura 5.3 encontra-se representado um processo autorregressivo e de médias móveis, ARMA(2,2), e as respetivas FAC e FACP empíricas.

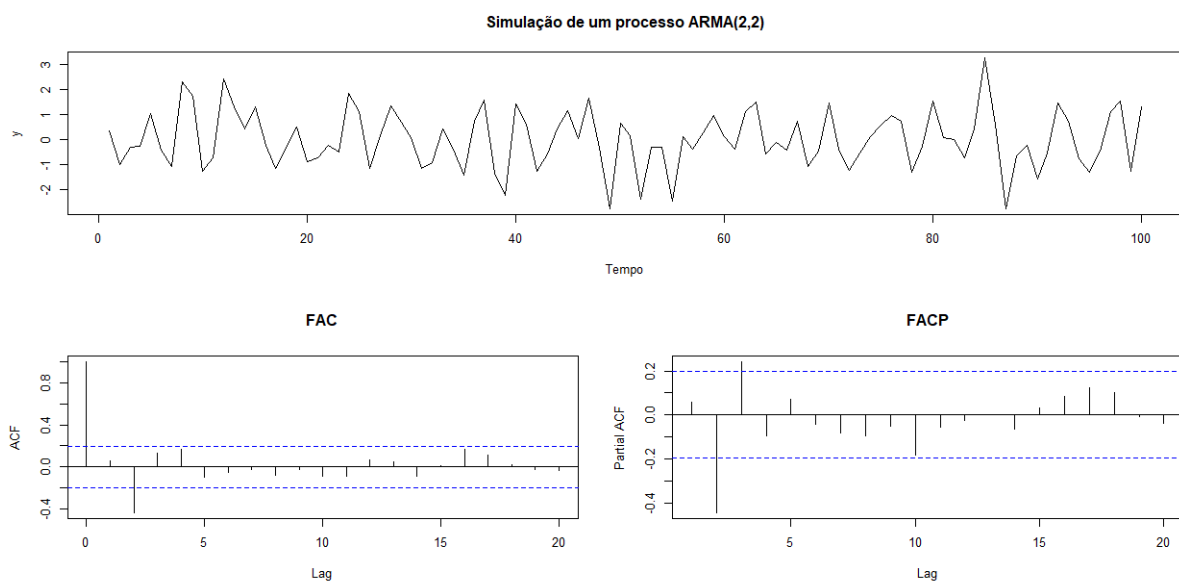


Figura 5.3: Simulação de um processo autorregressivo e de médias móveis e as respetivas FAC e FACP empíricas.

5.3 Modelos de Processos Não Estacionários

Numa perspetiva prática, a maioria das séries temporais é não estacionária. Quando tal se verifica, é necessária a remoção, nos dados, das fontes de variação não estacionárias como, por exemplo, a tendência e a sazonalidade, de forma a possibilitar o ajustamento de um modelo estacionário. Como foi referido anteriormente, se a série temporal observada for não estacionária na média, podem ser aplicadas uma ou várias diferenciações à mesma. Dessa forma, se se substituir Y_t por $\nabla^d Y_t$ na equação (5.7), obtém-se um modelo capaz de descrever séries estacionárias, ou seja, um modelo ARIMA. Este tipo de modelo é designado como modelo “integrado”, visto que o modelo estacionário que é ajustado aos dados diferenciados deve ser somado ou “integrado” de forma a devolver um modelo para os dados não estacionários. De notar que estes modelos, à semelhança dos modelos ARMA, podem ser generalizados para incluir termos sazonais, dando origem aos modelos SARIMA.

5.3.1 Processo Autorregressivo Integrado de Médias Móveis (ARIMA)

O processo Y_t diz-se um processo autorregressivo e de médias móveis integrado, $ARIMA(p, d, q)$, quando assume a expressão

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \epsilon_t \quad (5.8)$$

ou

$$\Phi_p(B) \nabla^d Y_t = \Theta_q(B) \epsilon_t, \quad (5.9)$$

onde $\nabla^d Y_t = (1 - B)^d Y_t$, com $d \geq 1$, é a série estacionária depois de diferenciada d vezes, $\phi_1, \phi_2, \dots, \phi_p$ são os parâmetros autorregressivos, $\theta_1, \theta_2, \dots, \theta_q$ são os parâmetros de médias móveis e $\Phi_p(B)$ e $\Theta_q(B)$ são os polinómios autorregressivos e de médias móveis regulares.

Dado que se trata de um processo não estacionário, um processo deste tipo apresenta uma FAC com coeficientes positivos e decaimento muito lento para zero, pelo que através desta se identifica que é necessário aplicar uma diferenciação. Na Figura 5.4 encontra-se representado um processo autorregressivo e de médias móveis integrado, $ARIMA(2, 1, 1)$, e as respetivas FAC e FACP empíricas.

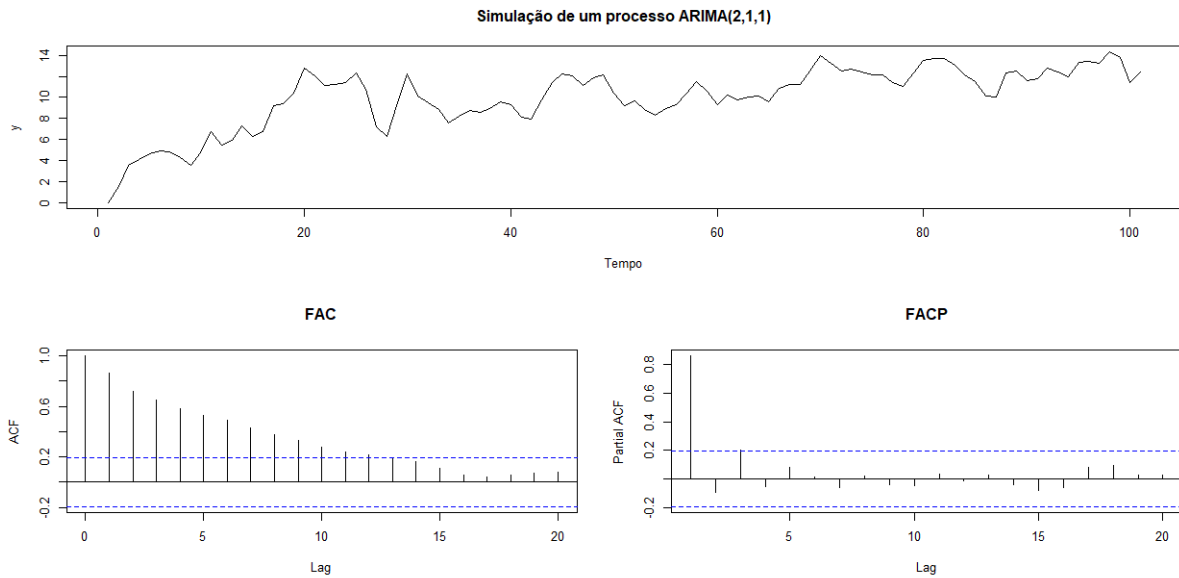


Figura 5.4: Simulação de um processo autorregressivo e de médias móveis integrado, ARIMA(2, 1, 1) e respectivas FAC e FACP empíricas.

5.3.2 Processo Autorregressivo Integrado de Médias Móveis Sazonal (SARIMA)

Nas séries temporais sazonais é previsível que a componente sazonal esteja de alguma forma relacionada com as componentes não sazonais. Isto significa que, se as observações vizinhas de uma série, $Y_t, Y_{t-1}, Y_{t-2}, \dots$, estão relacionadas, é muito provável que as observações vizinhas espaçadas em s unidades temporais, $Y_t, Y_{t-s}, Y_{t-2s}, \dots$, também estejam relacionadas. Por isso, pode estender-se o processo ARIMA(p, d, q) a um processo multiplicativo integrado sazonal, que generaliza todos os processos apresentados anteriormente e se representa por SARIMA(p, d, q)(P, D, Q) $_s$.

Antes de mais é conveniente apresentar a noção de diferenciação sazonal. Uma diferenciação sazonal não é mais do que a diferença entre a observação no instante t e a observação que ocorre s momentos antes, ou seja, a observação no instante $t - s$. Assim, quando uma série apresenta um comportamento periódico repetitivo, pode-se aplicar uma diferenciação sazonal, isto é,

$$\nabla_s Y_t = Y_t - Y_{t-s} = (1 - B^s)Y_t. \quad (5.10)$$

Como consequência, a série resultante desta diferenciação corresponde à mudança entre observações separadas por períodos de tempo s . Por exemplo, para uma série mensal, com $s = 12$, a série resultante de uma diferenciação sazonal representa a mudança que ocorre de ano para ano.

De uma forma semelhante ao que acontece para a diferenciação (regular) apresentada anteriormente, a diferenciação sazonal pode ser aplicada a uma série D vezes, originando o operador de diferenciação sazonal de ordem D , para qualquer inteiro $D \geq 1$, que se define por

$$\nabla_s^D Y_t = (1 - B^s)^D Y_t. \quad (5.11)$$

Deste modo, um processo Y_t diz-se um processo autorregressivo e de médias móveis integrado sazonal, SARIMA(p, d, q)(P, D, Q) $_s$, quando satisfaz a equação

$$\Phi_p(B)N_P(B^s)\nabla^d\nabla_s^D Y_t = \Theta_q(B)H_Q(B^s)\epsilon_t, \quad (5.12)$$

em que $\Phi_p(B)$, $N_P(B^s)$, $\Theta_q(B)$ e $H_Q(B^s)$ são os polinómios já referidos, d e D são as ordens de diferenciação das partes regular e sazonal, respetivamente. Normalmente, a necessidade de uma diferenciação sazonal pode ser indicada através da FAC de um processo, nomeadamente, quando esta decai lentamente nos *lags* múltiplos de s e é insignificante nos restantes (Shumway & Stoffer, 2011). Na Figura 5.5 encontra-se representado um processo SARIMA(2, 1, 1)(1, 1, 1) $_{12}$ e as respetivas FAC e FACP empíricas.

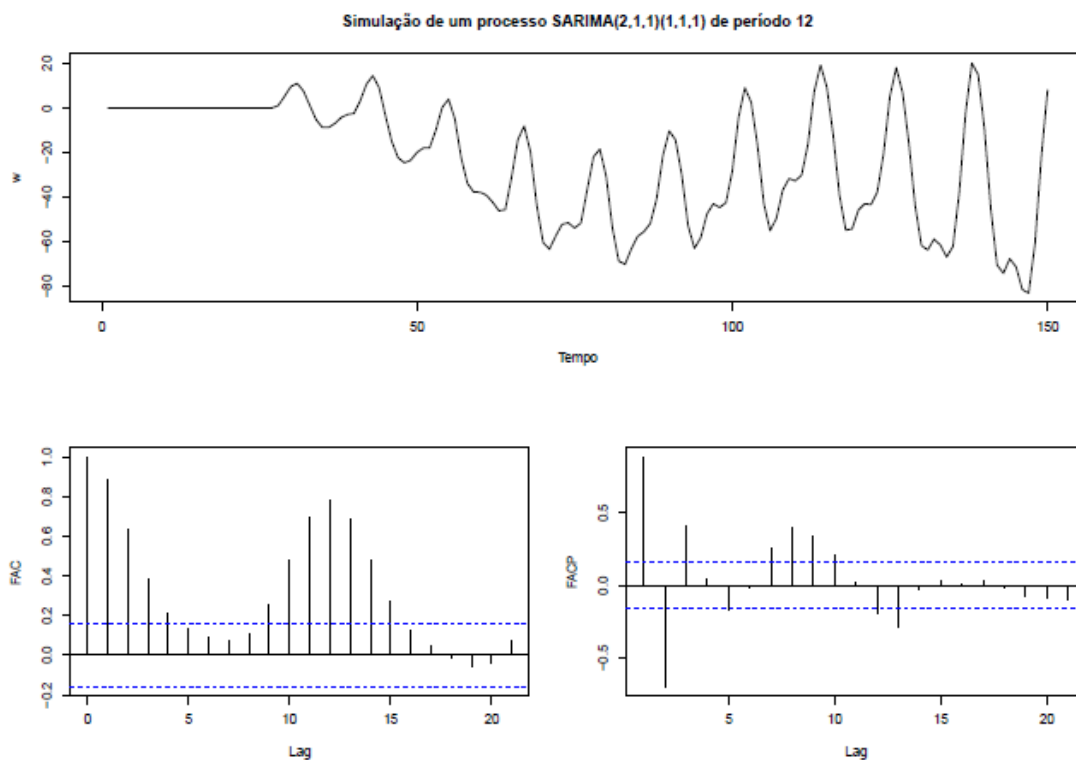


Figura 5.5: Simulação de um processo autorregressivo e de médias móveis integrado sazonal, SARIMA(2, 1, 1)(1, 1, 1) $_{12}$ e respetivas FAC e FACP empíricas.

5.4 Etapas da Metodologia Box-Jenkins

Nesta secção são explicadas com mais detalhe as etapas da metodologia Box-Jenkins, isto é, as etapas de identificação, estimação e diagnóstico, no caso da escolha de um modelo ARIMA.

A primeira etapa consiste na identificação de um modelo $ARIMA(p, d, q)$ que descreva a relação existente entre as suas observações. Esta etapa compreende três importantes passos na tentativa de identificação dos valores inteiros d, p e q .

1. Representação gráfica da série e respetiva estacionarização

O estudo de uma série temporal deve iniciar-se pela análise detalhada da sua representação gráfica, de modo a identificar a existência ou não de fontes de não estacionariedade. Caso a série em estudo não seja estacionária, deve-se proceder à sua estacionarização através de uma transformação adequada: com vista à estabilização da variância recorrer a transformações Box-Cox; para a eliminação da tendência utiliza-se a diferenciação regular. De notar que, caso a estabilização da variância seja necessária, esta deve ser efetuada antes de qualquer outra transformação.

2. Estimação das FAC e FACP da série original

Analisa-se o comportamento das FAC e FACP da série original, uma vez que as conclusões retiradas podem ser úteis para complementar à informação obtida através da representação gráfica (passo 1), nomeadamente no que diz respeito à utilização de diferenciações.

3. Estimação das FAC e FACP da série estacionária e identificação dos inteiros p e q

Identificada a ordens de diferenciação, d , a escolha do modelo que descreve a série temporal só é considerada concluída quando são determinados os inteiros p e q . Para identificar esses inteiros é efetuada a comparação do comportamento das FAC e FACP empíricas com o das FAC e FACP teórica.

Depois de identificados os modelos candidatos a descrever a série em estudo, segue-se a etapa de estimação dos seus parâmetros. Nesta fase é necessário recorrer a um *software* estatístico adequado, uma vez que esta etapa requer a aplicação de um conjunto de métodos numéricos e de cálculos computacionais com alguma complexidade. As principais abordagens utilizadas para ajustar os modelos dizem respeito a métodos de estimação por máxima verosimilhança ou a métodos de mínimos

quadrados. O método de máxima verosimilhança baseia-se na ideia de determinar os valores dos parâmetros que tornam mais verosímil a ocorrência de um conjunto de observações idênticas aquelas de que efetivamente se dispõe. Este método obtém estimativas dos parâmetros através de um processo iterativo em que se maximiza a função de verosimilhança dos estimadores (Box et al., 2016). Já o método dos mínimos quadrados é, possivelmente, o método estatístico mais utilizado na estimação de modelos. Na Tabela 5.1 são apresentados os comportamentos das FAC e FACP dos modelos de previsão, que permitem identificar os parâmetros p e q do modelo.

Tabela 5.1: Padrões teóricos das FAC e FACP dos modelos de previsão em séries temporais.

Processo	FAC	FACP
$AR(p)$	Decresce exponencialmente para zero com ou sem oscilações	Anula-se a partir de um intervalo maior do que p
$MA(q)$	Anula-se a partir de um intervalo maior do que q	Decresce exponencialmente para zero com ou sem oscilações
$ARMA(p, q)$	Decresce exponencialmente para zero com ou sem oscilações	Decresce exponencialmente para zero com ou sem oscilações

Depois de identificado o modelo ARIMA e estimados os respectivos parâmetros, é necessário verificar a adequação do modelo, ou seja, a última fase do processo. A fase de diagnóstico engloba duas etapas importantes: a avaliação da qualidade das estimativas obtidas e a avaliação da qualidade do ajustamento do modelo às observações da série em estudo.

Na primeira etapa é crucial analisar a significância estatística dos parâmetros estimados. Para tal, a cada parâmetro, diga-se β_i (onde $i = 1, \dots, m$ e m representa o número de parâmetros estimados), deve ser aplicado um teste de hipóteses apropriado que avalie a necessidade (ou não) de incluir esse parâmetro no modelo. Assim, interessa testar a hipótese de que β_i é estatisticamente nulo, isto é, $H_0 : \beta_i = 0$. A rejeição desta hipótese acontece, a um nível de significância α , quando a estatística T associada ao coeficiente estimado for, em valor absoluto, superior ou igual ao quantil $1 - \frac{\alpha}{2}$ de uma distribuição t de Student com $n - m$ graus de liberdade, (onde n representa o número de observações), isto é,

$$|T| = \left| \frac{\hat{\beta}_i}{\hat{\sigma}_{\hat{\beta}_i}} \right| \geq t_{1-\alpha/2; n-m}. \quad (5.13)$$

De forma equivalente, a hipótese nula é rejeitada, a um nível de significância α , quando o valor de

prova obtido é inferior ou igual a α . Segundo Caiado (2022), tendo sempre presente o princípio da parcimónia, devem incluir-se no modelo apenas os parâmetros que possam considerar-se significativamente diferentes de zero.

Relativamente à segunda etapa, ou seja, a avaliação da qualidade do ajustamento do modelo ARIMA, deve ser realizada uma análise do comportamento dos respetivos resíduos. Caso os resíduos apresentem um comportamento análogo ao de um ruído branco, pode afirmar-se que o modelo estimado descreve bem a série em estudo.

No caso de existir um modelo que não satisfaça os critérios aplicados em alguma das duas etapas, este deve ser rejeitado.

5.5 Análise dos Resíduos

Após efetuada a modelação de uma série temporal, os resíduos, que correspondem à informação não captada pelo modelo, podem ser calculados através da diferença entre os valores observados e os valores estimados correspondentes, isto é,

$$e_t = Y_t - \hat{Y}_t. \quad (5.14)$$

Como foi referido anteriormente, um bom modelo deve gerar resíduos com um comportamento análogo ao de um ruído branco, ou seja, estes devem apresentar média nula e devem satisfazer o pressuposto de não correlação. Adicionalmente, para a construção de intervalos de previsão, torna-se pertinente verificar se os resíduos têm variância constante e apresentam uma distribuição aproximadamente Normal (Hyndman & Athanasopoulos, 2021).

5.5.1 Distribuição Normal

A condição de normalidade pode ser avaliada tanto por análise gráfica, como por testes estatísticos ou, idealmente, por ambos. No caso das representações gráficas, as mais utilizadas são o histograma e o *QQ-plot*. Para indicar a normalidade, o histograma deve apresentar um comportamento semelhante ao da função densidade de uma distribuição Normal. Relativamente ao *QQ-plot*, visto que se trata de uma representação gráfica dos quantis reais e dos teóricos, este deve apresentar um conjunto de pontos que se posicione mais ou menos sobre uma reta correspondente à bissetriz dos quadrantes ímpares, ($y = x$).

Para uma verificação mais rigorosa, os testes estatísticos mais usuais são o teste de Shapiro-Wilk (para amostras de pequenas dimensões, geralmente com menos de 50 observações) e o teste de Lilliefors. Em ambos os testes é testada a hipótese nula de que "os erros seguem uma distribuição Normal".

5.5.2 Não Correlação

No que diz respeito ao pressuposto de não correlação, este deve ser verificado tanto individualmente como de forma conjunta. No caso individual, as autocorrelações dos resíduos podem ser avaliadas a partir da observação da FAC, sendo que, se o modelo for apropriado, esta deve apresentar um comportamento semelhante ao da FAC de um ruído branco, isto é, com autocorrelações não significativamente diferentes de zero.

De forma a testar as autocorrelações de forma conjunta, pode recorrer-se a um teste de Portmanteau. Dentro desta classe de testes, um dos mais utilizados é o teste de Ljung-Box, cuja estatística de teste Q é definida por

$$Q = n(n + 2) \sum_{j=1}^h \frac{\hat{\rho}_j^2}{n - j} \quad (5.15)$$

e segue aproximadamente uma distribuição do Qui-Quadrado com $h - m$ graus de liberdade (com h que corresponde ao número de autocorrelações a serem testadas e m ao número de parâmetros estimados). No caso de se rejeitar a hipótese nula, $H_0 : \rho_1 = \rho_2 = \dots = \rho_h = 0$, conclui-se que o modelo escolhido não é apropriado. Em relação ao valor de h , não existe um critério específico para a sua escolha, pelo que se apreende que a melhor abordagem passa por realizar o teste para vários valores distintos.

5.5.3 Média Nula e Variância Constante

A condição imposta à média dos resíduos pode ser facilmente analisada através do teste t para o valor médio. Neste teste rejeita-se a hipótese de média nula se

$$\left| \frac{\bar{e}\sqrt{n}}{s_e} \right| \geq t_{1-\frac{\alpha}{2};n-1}, \quad (5.16)$$

onde \bar{e} corresponde à média dos resíduos, n à dimensão da amostra, s_e ao desvio padrão dos resíduos e $t_{1-\frac{\alpha}{2};n-1}$ ao valor do quantil $1 - \frac{\alpha}{2}$ de uma distribuição t de Student com $n - 1$ graus de liberdade. De forma equivalente, a hipótese de média nula é rejeitada, a um nível de significância α , quando o valor de prova obtido é inferior ou igual a α . De notar que, este teste só deve ser aplicado quando os pressupostos

de normalidade e não correlação (independência, no caso de não se rejeitar a normalidade) se verificam. No caso da estabilidade da variância ou homocedasticidade, esta pode ser avaliada visualmente, através da análise do gráfico dos resíduos ao longo do tempo.

5.6 Seleção dos Modelos

Na modelação de uma série temporal pode existir mais do que um modelo que verifique os diferentes critérios de avaliação do diagnóstico especificados anteriormente, tornando a tarefa de escolher o melhor modelo mais complicada. Posto isto, devem procurar-se critérios de seleção de modelos que ponderem as estatísticas baseadas nos resíduos do modelo ajustado.

Um critério admissível para escolher o melhor modelo ARIMA poderia ser escolher o modelo que fornece a menor soma dos quadrados dos erros (ou erro quadrático médio) ou o maior valor para a função de verosimilhança. Contudo, esta abordagem nem sempre funciona, uma vez que, frequentemente, o erro quadrático médio pode ser reduzido e a função de verosimilhança aumentada simplesmente pelo aumento do número de parâmetros no modelo. De modo a contornar este problema, a função de verosimilhança deve ser penalizada por cada parâmetro adicional no modelo, ou seja, se o parâmetro extra não melhorar o valor da função de verosimilhança mais do que o valor da penalização, esse parâmetro não deve ser acrescentado ao modelo (Makridakis et al., 1998).

São diversos os critérios, baseados na função de verosimilhança, porém os mais utilizados são o critério de informação de Akaike (AIC) e o critério de informação Bayesiano (BIC, *Bayesian Information Criterion*).

5.6.1 Critério de Informação Akaike

Considere-se que um modelo com $m = p + q$ parâmetros foi ajustado a uma série com n observações. Akaike (1974), com objetivo de avaliar a qualidade do ajustamento, introduziu um critério baseado na quantidade de informação, definido por

$$AIC = -2 \log L + 2m, \quad (5.17)$$

onde L é a função de verosimilhança. De notar que nem todos os *softwares* estatísticos possuem a capacidade de determinar o AIC ou a função de verosimilhança L e, por essa razão, nem sempre é possível encontrar o AIC exato para um determinado modelo. Porém, uma aproximação útil para o AIC é

obtida através a aproximação

$$-2 \log L \approx n(1 + \log 2\pi) + n \log \sigma^2, \quad (5.18)$$

onde σ^2 representa a variância dos resíduos. Esta variância é facilmente estimada por qualquer *software* estatístico, permitindo, assim, que o AIC possa ser encontrado aproximadamente através da fórmula

$$AIC \approx n(1 + \log 2\pi) + n \log \hat{\sigma}^2 + 2m. \quad (5.19)$$

Por vezes, o primeiro termo em 5.19 é omitido por ser igual para todos os modelos.

O AIC não tem muito significado por si só e, por isso, só é útil em comparação com o AIC de outro modelo ajustado ao mesmo conjunto de dados. Desta forma deve escolher-se o modelo que tenha o menor AIC, tendo em consideração que uma diferença de valores de duas unidades (2) ou menos não é substancial. Nesses casos, deve optar-se pelo modelo mais simples, seja pela parcimónia, ou para obter um melhor ajustamento do modelo.

5.6.2 Critério de Informação Bayesiano

O critério de informação Bayesiano, proposto por Schwarz (1978), é definido por

$$BIC = -2 \log L + m \log(n), \quad (5.20)$$

onde L é a função de verosimilhança, m é o número de parâmetros do modelo e n é a dimensão da amostra.

Contrariamente ao AIC, o BIC depende da dimensão da amostra (n) pelo que, para $\log(n) > 2$, isto é, para uma amostra de dimensão superior a 7, a penalização do BIC é superior à penalização do AIC. Por consequência, a minimização do BIC leva, em geral, à seleção de modelos com um menor número de parâmetros do que os obtidos pela minimização do critério AIC, evitando, de certa forma, a sobrestimação do número de componentes.

5.7 Previsão Pontual e Intervalar

Após a escolha do modelo que melhor descreve a série temporal, podem ser calculadas as previsões, sejam estas pontuais ou intervalares. As previsões pontuais podem ser facilmente determinadas através da própria expressão do modelo escolhido. Efetivamente, de modo a obter previsões a h -passos, isto

é, para um instante $t + h$, basta calcular a esperança condicionada aos valores observados, ou seja, $E[Y_{t+h}|Y_1, Y_2, \dots, Y_t]$. Iniciando o processo para uma previsão a 1-passo, isto é, para $h = 1$, e repetindo-o para $h = 2, 3, \dots$, é, então, possível obter todas as previsões pretendidas.

Os intervalos de previsão usuais são construídos com base em estimativas do desvio padrão das próprias previsões. Partindo, assim, do pressuposto que os erros são independentes e que seguem uma distribuição Normal, a previsão intervalar para o instante $t + h$ é dada por

$$\left(\hat{y}_{t+h|t} - z_{1-\frac{\alpha}{2}} \hat{\sigma}_h, \hat{y}_{t+h|t} + z_{1-\frac{\alpha}{2}} \hat{\sigma}_h \right), \quad (5.21)$$

onde z é o quantil da distribuição Normal padrão, $1 - \alpha$ corresponde ao nível de confiança do intervalo e $\hat{\sigma}_h$ é a estimativa do desvio padrão da previsão para o passo h . Os níveis de confiança mais utilizados em intervalos de previsão são 90%, 95% e 99%, aos quais correspondem valores de z iguais a 1,645, 1,960 e 2,576, respetivamente. Geralmente, os intervalos de previsão aumentam conforme o horizonte de previsão, h , aumenta.

5.8 Métodos de Alisamento Exponencial

Os métodos de alisamento exponencial foram propostos por Holt (1957), Brown (1959) e Winters (1960) no final da década de 1960. Ao contrário dos modelos ARIMA descritos anteriormente, os métodos de alisamento exponencial são métodos não paramétricos. Estes são também utilizados para fazer previsão, porém são considerados métodos de previsão adaptativos, uma vez que são capazes de acompanhar mudanças no nível, declive e/ou padrão sazonal. Por esta razão são considerados mais eficientes quando são aplicados a séries cujas componentes, tendência e sazonalidade, variam com o tempo.

As previsões produzidas por estes métodos são consideradas médias ponderadas de observações anteriores, com os pesos a decair exponencialmente à medida que as observações se tornam mais distantes (Hyndman & Athanasopoulos, 2021). Por outras palavras, quanto mais recente é a observação, maior o peso que lhe será atribuído. Estes métodos têm vindo a destacar-se devido à sua versatilidade, rapidez, fiabilidade e facilidade de automatização (Armstrong, 2001). Em comparação com outros métodos de previsão, estes são mais simples de utilizar, apresentam um baixo custo de computação e uma facilidade de implementação em situações reais onde é necessário realizar previsões de várias séries simultaneamente. Porém, os métodos de alisamento exponencial também apresentam

algumas desvantagens como, por exemplo, o facto de estes serem métodos não paramétricos, o que impossibilita a realização de inferências estatísticas de uma forma direta. Por esta razão, é necessário ter em atenção a construção e interpretação de intervalos de previsão quando os erros de previsão não apresentarem um comportamento próximo de uma distribuição Normal de média zero.

De acordo com as diferentes formas que a tendência e a sazonalidade podem assumir (nula, aditiva ou multiplicativa), Pegels (1969) constrói uma primeira classificação dos métodos de alisamento exponencial. Posteriormente, e como consequência dos trabalhos de Gardner & Mckenzie (1985), Hyndman et al. (2002) e Taylor (2003), surgem novos métodos de alisamento exponencial, que levam ao alargamento da classificação inicial proposta por Pegels. O aparecimento de uma componente “amortecida” resulta na introdução de dois novos tipos de tendência, que culminam numa nova classificação dos métodos (Tabela 5.2).

Tabela 5.2: Classificação dos métodos de alisamento exponencial.

Componente Tendência	Componente Sazonalidade		
	N (Nenhuma)	A (Aditiva)	M (Multiplicativa)
N (Nenhuma)	N,N	N,A	N,M
A (Aditiva)	A,N	A,A	A,M
Aa (Aditiva amortecida)	Aa,N	Aa,A	Aa,M
M (Multiplicativa)	M,N	M,A	M,M
Ma (Multiplicativa amortecida)	Ma,N	Ma,A	Ma,M

Alguns dos métodos apresentados na Tabela 5.2 são conhecidos por outras designações, como por exemplo o método da célula (N, N) é denominado por método de alisamento exponencial simples, a célula (A, N) refere-se ao método linear de Holt e a célula (A_a, N) ao método amortecido. Outro método presente na Tabela 5.2 é o método aditivo de Holt-Winters na célula (A, A) e o método multiplicativo de Holt-Winters na célula (A, M) . As restantes entradas da Tabela correspondem a métodos que são pouco utilizados para previsão.

Na implementação dos métodos de alisamento exponencial são utilizadas várias notações que estão apresentadas na Tabela 5.3.

Tabela 5.3: Notações Utilizadas nos Métodos de Alisamento Exponencial.

Símb.	Definição	Símb.	Definição
\hat{Y}_t	Previsão a 1-passo para o instante t	h	Horizonte de previsão
α	Constante de alisamento associada ao nível	l_t	Nível no instante t
β	Constante de alisamento associada ao declive	b_t	Declive no instante t
γ	Constante de alisamento associada à sazonalidade	s_t	Fator sazonal no instante t
ϕ	Parâmetro de amortecimento da tendência	s	Período sazonal

5.8.1 Método de Alisamento Exponencial Simples

Como foi referido anteriormente, o método mais simples de todos é o método de alisamento exponencial simples (*Simple Exponential Smoothing* (SES)), desenvolvido por Brown (1959). Este método é apropriado para séries que não apresentem nem tendência nem padrão sazonal. Neste método considera-se que a previsão do período seguinte, \hat{Y}_{t+1} , é a previsão para o período corrente, \hat{Y}_t , ajustada através do respetivo erro de previsão, $e_t = Y_t - \hat{Y}_t$, isto é,

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t), \quad (5.22)$$

onde α é o parâmetro de alisamento que determina o peso dado a cada componente, e este varia entre 0 e 1. Para valores de α próximos de 0, a previsão de \hat{Y}_{t+1} sofre um ajustamento muito reduzido e é dado mais peso às observações mais distantes, ao invés de quando α toma valores próximos de 1, nesse caso sofre um ajustamento substancial e é dado mais peso às observações mais recentes.

A previsão de \hat{Y}_{t+1} também pode ser escrita como

$$\hat{Y}_{t+1} = \alpha Y_t + (1 - \alpha)\hat{Y}_t, \quad (5.23)$$

sendo considerada uma média pesada da previsão mais recente e da observação mais recente, uma vez que se baseia no peso α dado a Y_t e no peso $(1 - \alpha)$ dado a \hat{Y}_t . O valor de α pode ser utilizado para ajustar a sensibilidade do estimador a alterações no nível da série, quanto maior for o valor de α , maior é o peso atribuído a Y_t e mais sensível é a previsão às mudanças no nível (Wang, 2006).

Se se substituir na equação 5.23 o valor de \hat{Y}_t por $\alpha Y_{t-1} + (1 - \alpha)\hat{Y}_{t-1}$ obtém-se

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1 - \alpha)Y_{t-1} + (1 - \alpha)^2 \hat{Y}_{t-1},$$

e, repetindo esse processo sucessivamente, tem-se

$$\hat{Y}_{t+1} = (1 - \alpha)^t \hat{Y}_1 + \alpha \sum_{i=0}^{t-1} (1 - \alpha)^i Y_{t-i},$$

onde se comprova que \hat{Y}_{t+1} representa uma média móvel pesada de todas as observações anteriores, com pesos exponencialmente decrescentes (Hyndman & Athanasopoulos, 2021).

Existe ainda outra forma de representar a equação 5.22, utilizando a componente nível, l_t . Nesta considera-se que o nível da série (l_t) é igual à previsão \hat{Y}_{t+1} , e portanto

$$l_t = \alpha Y_t + (1 - \alpha)l_{t-1} \hat{Y}_{t+h} = l_t. \quad (5.24)$$

Uma das principais limitações dos métodos recursivos recai sobre a inicialização do processo. Efetivamente, se uma nova previsão depende sempre de uma previsão anterior, é fundamental iniciar o processo com um valor pré-definido para a primeira previsão. Posto isto, Makridakis et al. (1998) propõem a utilização da primeira observação como a previsão inicial, ou seja, $\hat{Y}_1 = Y_1$. Outra possibilidade é utilizar a média dos primeiros quatro ou cinco valores da série temporal.

Relativamente à previsão para períodos futuros, a mais de um passo, constata-se que, uma vez que este método é adequado apenas para séries sem tendência, sazonalidade ou qualquer outro tipo de padrão subjacente, a função de previsão é “constante”, i.e., $\hat{Y}_{t+h} = \hat{Y}_{t+1}$, com $h = 2, 3, \dots$

5.8.2 Método Linear de Holt

Uma vez que o método de alisamento exponencial simples não pode ser aplicado a séries que apresentem tendência, Holt (1957) desenvolveu uma extensão do modelo anterior, adicionando uma equação de atualização para o declive. Assumindo b_t como o estimador do declive no tempo t , o método linear de Holt (ou apenas método de Holt) utiliza duas equações para realizar a atualização das componentes nível e declive

$$\textbf{Nível: } l_t = \alpha Y_t + (1 - \alpha)(l_{t-1} + b_{t-1}), \quad (5.25)$$

$$\textbf{Declive: } b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \quad (5.26)$$

onde l_t representa a estimação para o nível da série no período t , b_t é a estimação do declive da série no mesmo período t e α e β apresentam valores compreendidos entre 0 e 1.

Para realizar previsões a partir deste método é necessário adicionar um segundo termo, hb_t , uma vez que neste caso estamos perante séries com tendência. Sendo assim,

$$\textbf{Previsão: } \hat{Y}_{t+h} = l_t + hb_t, \quad h = 1, 2, \dots \quad (5.27)$$

Em relação à inicialização do processo, neste caso é necessário estimar l_1 e b_1 . Para esse fim, Makridakis et al. (1998) recomendam que se defina $\hat{l}_1 = Y_1$ e, no caso de b_1 existem duas possibilidades, $\hat{b}_1 = Y_2 - Y_1$ ou $\hat{b}_1 = (Y_4 - Y_1)/3$. Existe ainda outra alternativa que passa por utilizar a regressão e, através da aplicação do método dos mínimos quadrados, determinar \hat{l}_1 e \hat{b}_1 com base em alguns valores iniciais da série temporal.

À semelhança de Holt, Brown (1963) também propõe um método de previsão para séries com tendência denominado método de alisamento exponencial duplo. Ao contrário do método de Holt, este utiliza apenas uma constante de alisamento e aplica o método de alisamento exponencial duas vezes. A implementação do método de alisamento duplo de Brown é realizado através das seguintes equações de atualização

$$m_t = \alpha Y_t + (1 - \alpha)m_{t-1}, \quad (5.28)$$

$$d_t = \alpha m_t + (1 - \alpha)d_{t-1}, \quad (5.29)$$

onde m_t corresponde à série de alisamento exponencial simples, d_t à série de alisamento exponencial duplo e α à constante de alisamento.

As previsões são obtidas através da equação 5.27, utilizada no método de Holt, porém considerando novas estimativas para o nível e declive

$$\textbf{Nível: } l_t = 2m_t - d_t, \quad (5.30)$$

$$\textbf{Declive: } b_t = \frac{\alpha}{1 - \alpha}(m_t - d_t). \quad (5.31)$$

Ao realizar a comparação entre as equações 5.25 e 5.30, 5.26 e 5.31, constata-se que quando $\alpha = \beta$, o método de Holt e o método de alisamento exponencial duplo são equivalentes (Caiado, 2022).

Posto isto, pode-se concluir que o método de alisamento exponencial duplo é um caso particular do método de Holt.

5.8.3 Método de Holt-Winters

O método de Holt apresentado anteriormente era designado para séries que apresentassem somente tendência. Visto que algumas séries, para além de tendência, apresentam também sazonalidade, o método de Holt foi expandido por Winters (1960) para ser aplicado nessas mesmas séries. Este método foi intitulado como método de Holt-Winters e apresenta três equações de atualização, uma para o nível, outra para a tendência e outra para a sazonalidade. O método de Holt-Winters é constituído por dois métodos distintos, dependendo se a sazonalidade se encontra na sua forma aditiva ou multiplicativa. Se esta se apresentar na sua forma aditiva, o método adapta-se a séries com amplitude sazonal independente ao nível da série, por outro lado, se se apresentar na sua forma multiplicativa, o método é utilizado em séries cuja amplitude sazonal é proporcional ao nível da série.

Método Aditivo de Holt-Winters

O modelo de decomposição aditivo define-se por

$$Y_t = T_t + S_t + \epsilon_t. \quad (5.32)$$

O método aditivo de Holt-Winters adapta-se a séries que apresentam uma amplitude sazonal independente ao nível da série. Este é menos utilizado do que o método multiplicativo e é constituído pelas seguintes equações de atualização:

$$\textbf{Nível: } l_t = \alpha(Y_t - s_{t-s}) + (1 - \alpha)(l_{t-1} + b_{t-1}), \quad (5.33)$$

$$\textbf{Tendência: } b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}, \quad (5.34)$$

$$\textbf{Sazonalidade: } s_t = \gamma(Y_t - l_t) + (1 - \gamma)s_{t-s}, \quad (5.35)$$

onde $0 \leq \gamma \leq 1$.

As previsões são obtidas a partir da equação

$$\text{Previs\~{a}o: } \hat{Y}_{t+h} = l_t + hb_t + s_{t-s+h_s^+}, \quad h = 1, 2, \dots, \quad (5.36)$$

onde $h_s^+ = [(h - 1) \text{ mod } s] + 1$.

Para se iniciar este m\~{e}todo \u00e9 necess\u00e1rio definir os valores iniciais do n\u00edvel, declive e os \u00edndices de sazonalidade. Para tal utilizam-se as seguintes equa\u00e7\u00f5es:

$$\hat{l}_s = \frac{1}{s} \sum_{i=1}^s Y_i, \quad (5.37)$$

$$\hat{b}_s = \frac{1}{s^2} \left(\sum_{i=s+1}^{2s} Y_i - \sum_{i=1}^s Y_i \right), \quad (5.38)$$

$$\hat{s}_i = Y_i - \hat{l}_s, \text{ para } i = 1, \dots, s. \quad (5.39)$$

M\u00e9todo Multiplicativo de Holt-Winters

O modelo de decomposi\u00e7\u00e3o multiplicativo \u00e9 definido por

$$Y_t = T_t \times S_t + \epsilon_t. \quad (5.40)$$

Ao contr\u00e1rio do m\u00e9todo aditivo, o m\u00e9todo multiplicativo de Holt-Winters \u00e9 utilizado em s\u00e9ries cuja amplitude sazonal \u00e9 proporcional ao n\u00edvel da s\u00e9rie. Este m\u00e9todo \u00e9 tamb\u00e9m composto por tr\u00eas equa\u00e7\u00f5es de atualiza\u00e7\u00e3o

$$l_t = \alpha \frac{Y_t}{s_{t-s}} + (1 - \alpha)(l_{t-1} + b_{t-1}), \quad (5.41)$$

$$b_t = \beta (l_t + l_{t-1}) + (1 - \beta) b_{t-1}, \quad (5.42)$$

$$s_t = \gamma \frac{Y_t}{l_t} + (1 - \gamma) s_{t-s}, \quad (5.43)$$

e as previs\u00f5es s\u00e3o obtidas atrav\u00e9s da equa\u00e7\u00e3o

$$\hat{Y}_{t+h} = (l_t + hb_t) s_{t-s+h_s^+}, \quad h = 1, 2, \dots \quad (5.44)$$

Para a inicialização deste método é necessário definir os valores iniciais do nível, declive e para os índices de sazonalidade através das seguintes equações

$$\hat{l}_s = \frac{1}{s} \sum_{i=1}^s Y_i, \quad (5.45)$$

$$\hat{b}_s = \frac{1}{s^2} \left(\sum_{i=s+1}^{2s} Y_i - \sum_{i=1}^s Y_i \right), \quad (5.46)$$

$$\hat{s}_i = \frac{Y_i}{\hat{l}_s}, \text{ para } i = 1, \dots, s. \quad (5.47)$$

A escolha das constantes de alisamento α , β e γ é obtida através de métodos numéricos de otimização de maneira a que o erro quadrático médio das previsões a 1-passo seja mínimo.

5.8.4 Outros Métodos de Alisamento Exponencial

Ao longo do tempo foram surgindo novas propostas de métodos para realizar a previsão de séries temporais. De acordo com alguns estudos realizados, o método linear de Holt tende a sobrepor-se aos dados quando utilizado para previsão a longo prazo. Para contornar esse problema, Gardner e Mckenzie (1985) introduziram um novo parâmetro ϕ ao método de Holt.

No caso das séries não sazonais, é introduzido o método de Holt-Winters com tendência aditiva amortecida, referente à célula (A_a , N) da Tabela 5.2. Este método tem como equações de atualização

$$\textbf{Nível: } l_t = \alpha Y_t + (1 - \alpha)(l_{t-1} + \phi b_{t-1}), \quad (5.48)$$

$$\textbf{Declive: } b_t = \beta (l_t - l_{t-1}) + (1 - \beta) \phi b_{t-1}, \quad (5.49)$$

onde α , β e ϕ variam entre 0 e 1.

Já a previsão a h -passos é obtida por

$$\textbf{Previsão: } \hat{Y}_{t+h} = l_t + \sum_{i=1}^h \phi^i b_t, \quad h = 1, 2, \dots \quad (5.50)$$

De notar que quando $\phi = 1$, este método corresponde ao método linear de Holt.

No caso de séries sazonais, foram propostos dois métodos de tendência amortecida, o método de Holt-Winters multiplicativo com tendência aditiva amortecida e o método de Holt-Winters aditivo com

tendência aditiva amortecida, células (A_a, M) e (A_a, A) da Tabela 5.2, respectivamente. Constatase ainda que, quando $\phi = 1$, estes métodos são equivalentes aos métodos multiplicativo e aditivo de Holt-Winters, respectivamente.

Em relação ao método de Holt-Winters aditivo com tendência aditiva amortecida, este apresenta as seguintes equações de atualização e previsão

$$\mathbf{Nível:} \quad l_t = \alpha (Y_t - s_{t-s}) + (1 - \alpha)(l_{t-1} + \phi b_{t-1}), \quad (5.51)$$

$$\mathbf{Tendência:} \quad b_t = \beta (l_t - l_{t-1}) + (1 - \beta)\phi b_{t-1}, \quad (5.52)$$

$$\mathbf{Sazonalidade:} \quad s_t = \gamma (Y_t - l_t) + (1 - \gamma) s_{t-s}, \quad (5.53)$$

$$\mathbf{Previsão:} \quad \hat{Y}_{t+h} = l_t + \sum_{i=1}^h \phi^i b_t + s_{t-s+h_s^+}, \quad h = 1, 2, \dots, \quad (5.54)$$

onde $h_s^+ = [(h - 1) \text{ mod } s] + 1$.

Já o método de Holt-Winters multiplicativo com tendência aditiva amortecida é constituído pelas seguintes equações de atualização e previsão

$$\mathbf{Nível:} \quad l_t = \alpha \frac{Y_t}{s_{t-s}} + (1 - \alpha)(l_{t-1} + \phi b_{t-1}), \quad (5.55)$$

$$\mathbf{Tendência:} \quad b_t = \beta (l_t - l_{t-1}) + (1 - \beta) \phi b_{t-1}, \quad (5.56)$$

$$\mathbf{Sazonalidade:} \quad s_t = \gamma \frac{Y_t}{l_t} + (1 - \gamma) s_{t-s}, \quad (5.57)$$

$$\mathbf{Previsão:} \quad \hat{Y}_{t+h} = (l_t + \sum_{i=1}^h \phi^i b_t) s_{t-s+h_s^+}, \quad h = 1, 2, \dots, \quad (5.58)$$

onde $h_s^+ = [(h - 1) \text{ mod } s] + 1$.

5.8.5 Intervalos de Previsão

Após serem obtidas as previsões pontuais, torna-se útil construir um intervalo de previsão para cada instante temporal. Os intervalos de previsão são, normalmente, baseados na medida do erro quadrático médio (EQM),

$$\text{EQM} = \frac{1}{n} \sum_{t=1}^n e_t^2 = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2, \quad (5.59)$$

que constitui uma estimativa da variância do erro de previsão a 1-passo (i.e., com horizonte temporal de uma unidade temporal). Se se admitir a hipótese de que os erros de previsão têm distribuição aproximadamente Normal de média zero, pode, então, construir-se um intervalo de previsão aproximado para cada instante do tempo.

Posto isto, supondo a normalidade dos erros de previsão, segundo Makridakis et al. (1998), o intervalo de previsão a 1-passo para o instante t pode ser obtido através da expressão

$$\left(\hat{Y}_t - z_{1-\alpha/2} \sqrt{\text{EQM}}, \hat{Y}_t + z_{1-\alpha/2} \sqrt{\text{EQM}} \right),$$

onde z é o quantil da distribuição Normal padrão e $1 - \alpha$ corresponde ao nível de confiança do intervalo.

No caso do objetivo da previsão, no que se refere ao horizonte temporal, for de médio ao longo prazo, devem ser calculadas previsões a h -passos, utilizando apenas a informação disponível da série até ao momento t . Assim sendo, a medida do EQM é determinada com base nos erros de previsão a h -passos e este é definido como

$$\text{EQM}_{(h)} = \frac{1}{n-h} \sum_{t=h+1}^n (Y_t - \hat{Y}_{t|h})^2, \quad (5.60)$$

onde h é o horizonte temporal. No caso da série temporal apresentar sazonalidade, a expressão (5.60) deve ser substituída por

$$\text{EQM}_{(h)} = \frac{1}{n-h-s+1} \sum_{t=h+s}^n (Y_t - \hat{Y}_{t|h})^2,$$

em que s corresponde ao período sazonal.

Por fim, de modo a construir intervalos de previsão empíricos (ou aproximados) multi-passos (h unidades temporais) deve utilizar-se a expressão

$$\left(\hat{Y}_{t+h} - z_{1-\alpha/2} \sqrt{\text{EQM}_{(h)}}, \hat{Y}_{t+h} + z_{1-\alpha/2} \sqrt{\text{EQM}_{(h)}} \right). \quad (5.61)$$

5.9 Medidas de Avaliação

Nesta secção são apresentadas as medidas de avaliação utilizadas de modo a proceder à avaliação da qualidade preditiva das metodologias aplicadas: uma medida dependente da escala, EQM, e a sua correspondente na mesma escala dos dados, REQM; uma medida percentual, EPAM; uma medida escalada, EAM; e uma medida alternativa, estatística U de Theil.

O Erro Quadrático Médio (EQM), definido em 5.59, consiste no quadrado do valor médio dos desvios entre os valores observados e as previsões para os instantes $1, 2, \dots, n$. Esta medida é caracterizada pela dependência da escala dos dados e é calculada através do erro de previsão a 1-passo, $e_t = Y_t - \hat{Y}_t$. Regularmente, a raiz do erro quadrático médio, $REQM = \sqrt{EQM}$, é preferida ao EQM, uma vez que esta permite reduzir a grandeza dos valores para a mesma escala dos dados.

Ao realizar a comparação entre métodos de previsão, e visto que se $Y_t = \hat{Y}_t$ se obtém $EQM = 0$, admite-se que o método mais preciso é o que apresenta menor EQM e, conseqüentemente, menos REQM. De notar que estas medidas descritas são mais sensíveis à presença de *outliers* em comparação, por exemplo, ao Erro Absoluto Médio (EAM), que é definido por

$$EAM = \frac{1}{n} \sum_{t=1}^n |e_t| = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|.$$

O Erro Percentual Absoluto Médio (EPAM) consiste na percentagem média do erro de previsão em relação à grandeza das observações e, por esta razão, é considerada uma medida de erros percentuais, definida como

$$EPAM = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100 \quad (\%).$$

Tal como no EQM, note-se que se $Y_t = \hat{Y}_t$ tem-se $EPAM = 0\%$ logo, quanto menor for o EPAM mais preciso é o método de previsão.

Por último, a Estatística U de Theil que, à semelhança do EQM, os erros são elevados ao quadrado e, dessa forma, é dado maior peso aos grandes erros em detrimento dos mais pequenos. Esta estatística é definida como

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} (f_{t+1} - a_{t+1})^2}{\sum_{t=1}^{n-1} (a_{t+1})^2}}, \quad (5.62)$$

onde $f_{t+1} = (\hat{Y}_{t+1} - Y_t)/Y_t$ e $a_{t+1} = (Y_{t+1} - Y_t)/Y_t$, e estas correspondem às mudanças relativas prevista e real, respetivamente.

Substituindo na expressão (5.62) os valores de f_{t+1} e a_{t+1} obtém-se, numa versão simplificada,

$$U = \sqrt{\frac{\sum_{t=1}^{n-1} \left(\frac{\hat{Y}_{t+1} - Y_{t+1}}{Y_t} \right)^2}{\sum_{t=1}^{n-1} \left(\frac{Y_{t+1} - Y_t}{Y_t} \right)^2}}.$$

Analisando as expressões acima, constata-se que:

- $U = 0$ se $f_{t+1} = a_{t+1}$, para $t = 1, 2, \dots, n - 1$, i.e., quando as previsões são exatas;
- $U = 1$ se $f_{t+1} = 0$, i.e., quando os erros de previsão do método adotado são iguais aos erros obtidos pelo método de referência. Neste caso, pode dizer-se que o método adotado e o método de referência são igualmente precisos;
- $U > 1$ se f_{t+1} se encontra na direção oposta (sinais diferentes) de a_{t+1} , uma vez que o numerador será maior do que o denominador. Neste caso, o método de previsão adotado é menos preciso do que o método de referência e, portanto, não deve ser o escolhido;
- $U < 1$ se f_{t+1} se encontra na mesma direção (sinais iguais) de a_{t+1} . Assim, o método de previsão adotado é mais preciso do que o método de referência e, portanto, deve ser o escolhido. Quanto menor o valor de U , melhor é o método de previsão relativamente ao método de referência.

Tal como nas medidas de avaliação definidas anteriormente, considera-se que na comparação entre diferentes métodos de previsão, o mais preciso é o que apresenta o valor de U mais baixo.

6 Método de Lee-Carter e Modelo de Poisson para a População Portuguesa

A taxa de mortalidade é uma função fundamental para o cálculo das restantes funções presentes numa tábua de mortalidade. De maneira a ajustar a taxa e obter previsões desta para anos futuros é necessário utilizar métodos para esse fim, para isso são utilizados métodos como o Método de Lee-Carter e o Modelo de Poisson. Neste capítulo é apresentada uma análise exploratória dos dados da população portuguesa e são aplicados aos dados os métodos de ajustamento da taxa de mortalidade. Posteriormente, são construídas as tábuas de mortalidade para os anos em estudo e são também calculadas as taxas vitalícias necessárias para mais tarde serem calculadas as pensões vitalícias dos pensionistas.

6.1 Análise Descritiva dos Dados

Nesta secção são apresentados os dados utilizados da população portuguesa e é realizada uma breve análise dos mesmos em três casos: no sexo masculino, feminino e sem distinção de sexo. Todos os dados utilizados neste estudo referentes à população portuguesa foram retirados da *Human Mortality Database* (HMD). A HMD é o principal recurso mundial de dados científicos sobre mortalidade nos países desenvolvidos. Este fornece estimativas detalhadas e harmonizadas de alta qualidade sobre mortalidade e população a investigadores, estudantes, jornalistas, analistas, e outros interessados na longevidade humana (Database, 2022). Foram retiradas do *site*, <https://www.mortality.org/>, as bases de dados referentes ao número de mortes, $d_{x,t}$, e à exposição ao risco, $E_{x,t}$, durante os anos 1940-2020, com e sem distinção de sexo. Para analisar o comportamento de cada série temporal foram obtidas representações gráficas como gráficos de linhas e *boxplots*, para visualizar como se distribuem os dados e para verificar se estes apresentam valores discrepantes, isto é, *outliers*. Para além dos gráficos foram também calculadas algumas estatísticas descritivas de cada série temporal.

6.1.1 Caso 1: Sexo Masculino

Em relação ao número de mortes observadas ao longo do tempo, nas Figuras A.1, A.2, A.3 e A.4 do Apêndice A estão apresentadas as representações gráficas de todos os anos do estudo e, a título de exemplo, na Figura 6.6 estão exibidos os gráficos dos anos 1940, 1970, 2000 e 2020. Através desta observa-se que nos anos 1940 e 1970 o número de mortes é mais elevado em idades mais novas, isto é, entre os 0-5 anos, e entre os 50-80 anos apresenta valores elevados, mas não tanto como para as idades iniciais. Passado alguns anos, nomeadamente nos gráficos dos anos 2000 e 2020, é notório o decréscimo do número de mortes na faixa etária dos 0-5 anos e o aumento na faixa etária dos 50-90 anos, sendo que o valor mais elevado dentro deste intervalo de tempo foi registado em 2020 na idade 84 anos.

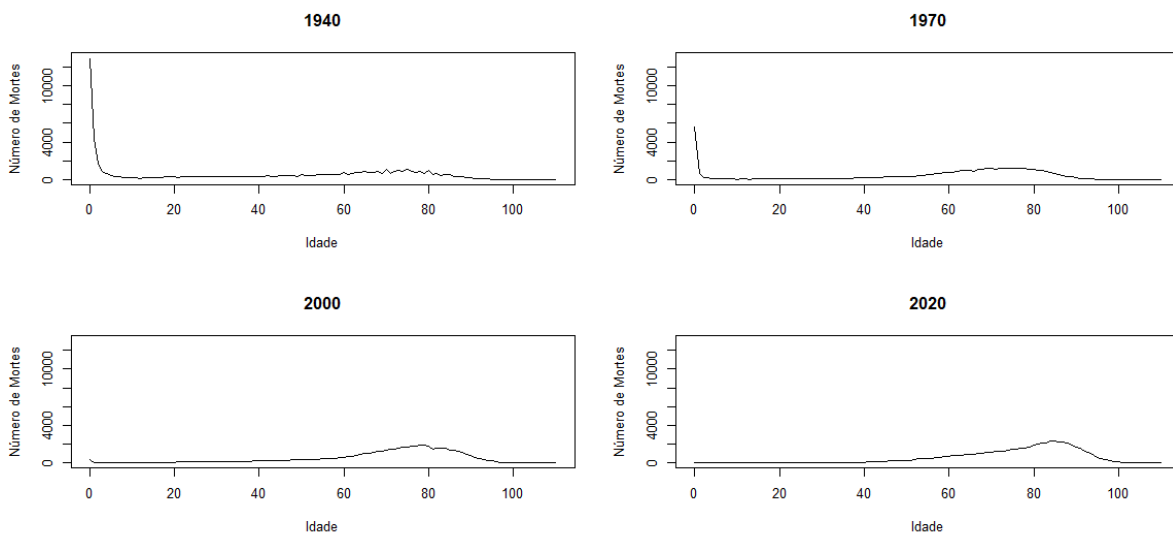


Figura 6.6: Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sexo Masculino.

No que respeita à presença de *outliers*, na Figura 6.7 constata-se que existem vários na maior parte dos anos. Porém, depois de analisados separadamente, constatou-se que estes valores não correspondiam a erros, mas sim a valores extremos comparativamente com os restantes, sendo que, por esta razão, foram utilizados para o estudo. Mais detalhadamente, nos anos 1940 a 1982 as observações consideradas como *outliers* correspondiam a idades entre os 0-2 anos, uma vez que nestes anos a mortalidade infantil à nascença era muito elevada devido à falta de condições da altura em questão, como, por exemplo, a alimentação, a habitação, o acesso aos cuidados médicos, a adequação e a qualidade dos serviços de saúde, entre outros (Barreto et al., 2014). A partir de 1983 até 2020 os

outliers correspondem a idades da faixa etária 75-86 anos. Estes valores explicam-se uma vez que nestas idades é normal existirem mais mortes devido à idade avançada ou até por doença. De notar que estes valores já eram elevados entre os anos 1940-1982, porém, devido ao decréscimo de mortes em idades iniciais (0-2 anos), estes valores apenas são considerados *outliers* entre 1983-2020 uma vez que são mais discrepantes nestes anos do que nos anteriores.

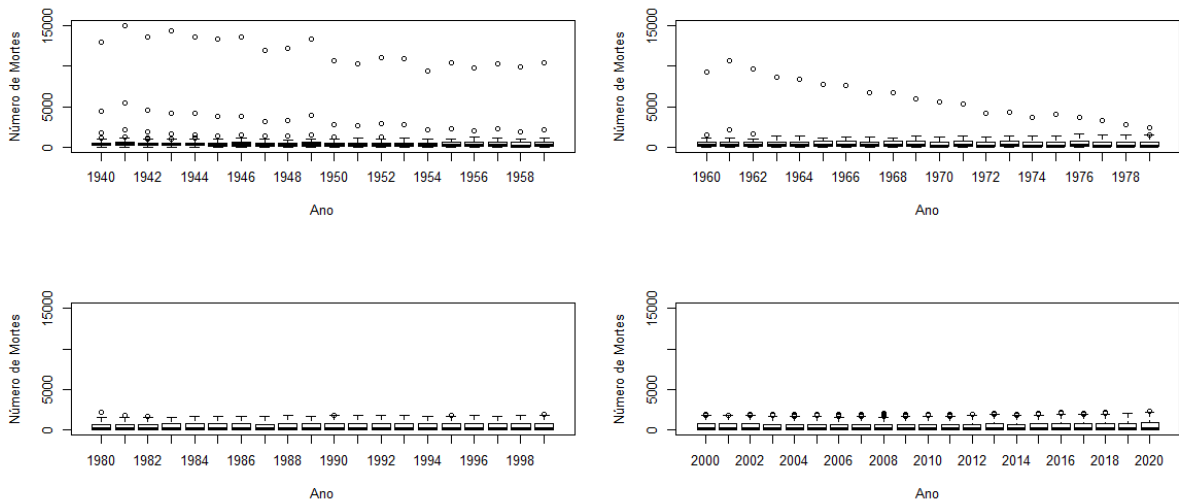


Figura 6.7: *Boxplots* do número de mortes nos anos 1940-2020 - Sexo Masculino.

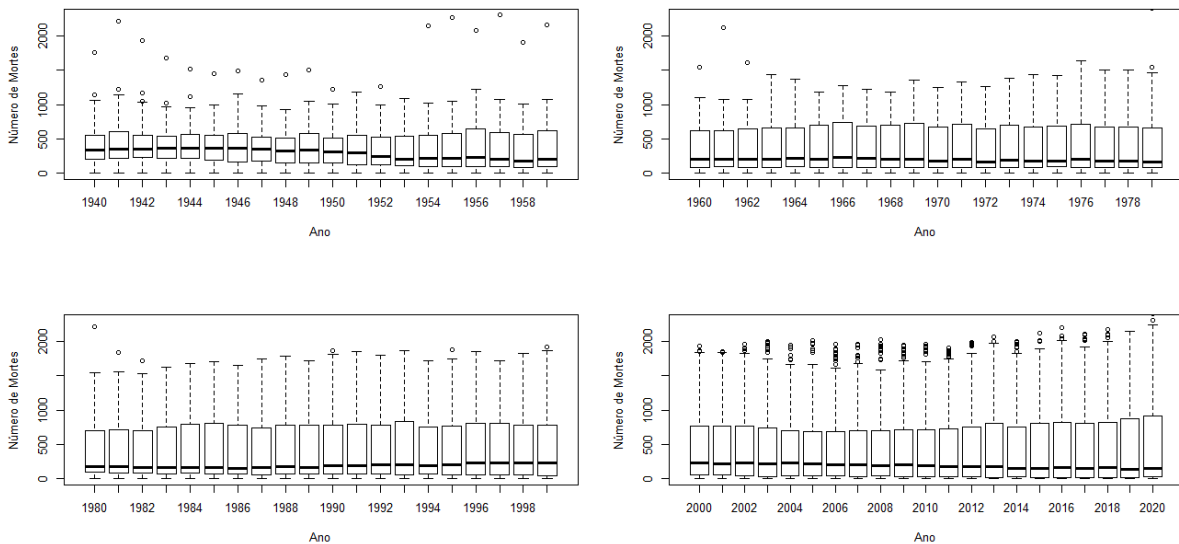


Figura 6.8: *Zoom* dos *boxplots* do número de mortes nos anos 1940-2020 - Sexo Masculino.

De forma a entender a dispersão dos dados, na Figura 6.8 é apresentado um *zoom* dos *boxplots* do

número de mortes e constata-se que todos os anos apresentam uma assimetria positiva, isto é, as caudas direitas são muito mais alongadas do que as esquerdas. Observa-se também que a dispersão dos dados tende a aumentar com o passar dos anos, atingindo o valor máximo em 2020.

Relativamente à exposição ao risco, nas Figuras A.5, A.6, A.7 e A.8 do Apêndice A estão apresentados os gráficos obtidos com os dados de todos os anos do estudo, porém, como exemplo, são apresentados na Figura 6.9 o número de expostos ao risco nos anos 1940, 1970, 2000 e 2020. Através desta observa-se que o comportamento dos dados é semelhante em 1940 e 1970, uma vez que decresce desde a idade 0 até 90 anos, sendo que a partir desse momento é quase constante. Nos anos seguintes, nomeadamente em 2000 e 2020, o maior valor deixa de ser na idade 0 anos e passa a ser nas faixas 20-30 e 40-50 anos, respetivamente.

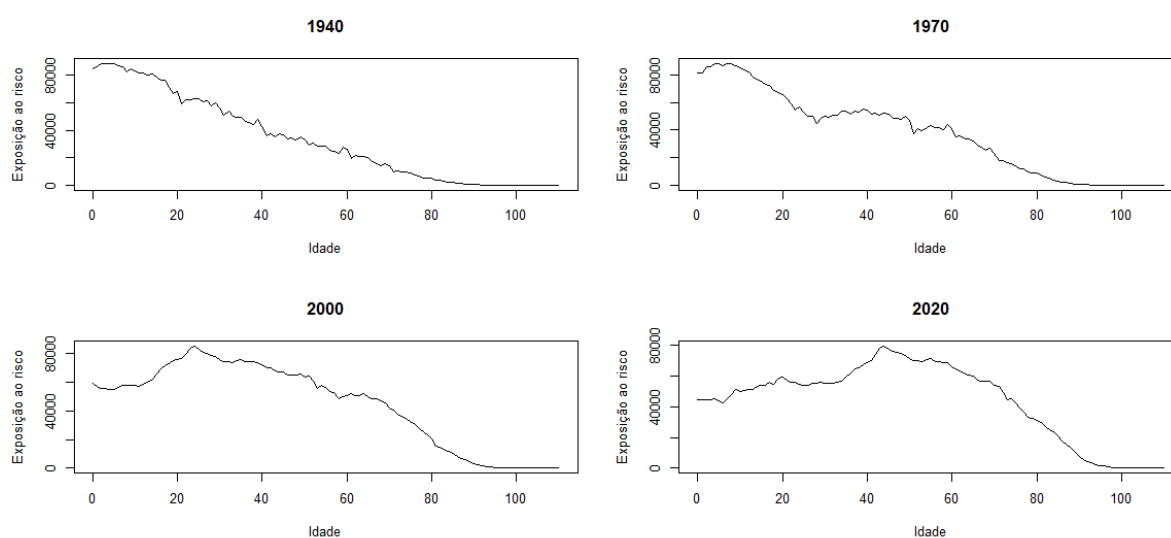


Figura 6.9: Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sexo Masculino.

Quanto à análise dos *boxplots* da exposição ao risco de cada ano, a partir da Figura 6.10 observa-se que nesta variável não existem *outliers*. Em relação à dispersão dos dados, o ano em que os dados são mais dispersos é 1962, que consequentemente é o ano onde se verifica o maior valor de exposição ao risco na idade 0 anos e, em relação à simetria, todos os anos apresentam uma assimetria negativa, isto é, a cauda esquerda é mais alongada do que a cauda direita, com a exceção dos dados dos anos 1940-1944 e 1948-1954 que são assimétricos positivos. Através da mediana constata-se que desde 1980 até 2020 os valores de exposição ao risco têm sido praticamente constantes e observa-se também que a dispersão dos dados tem vindo a diminuir com o passar dos anos.

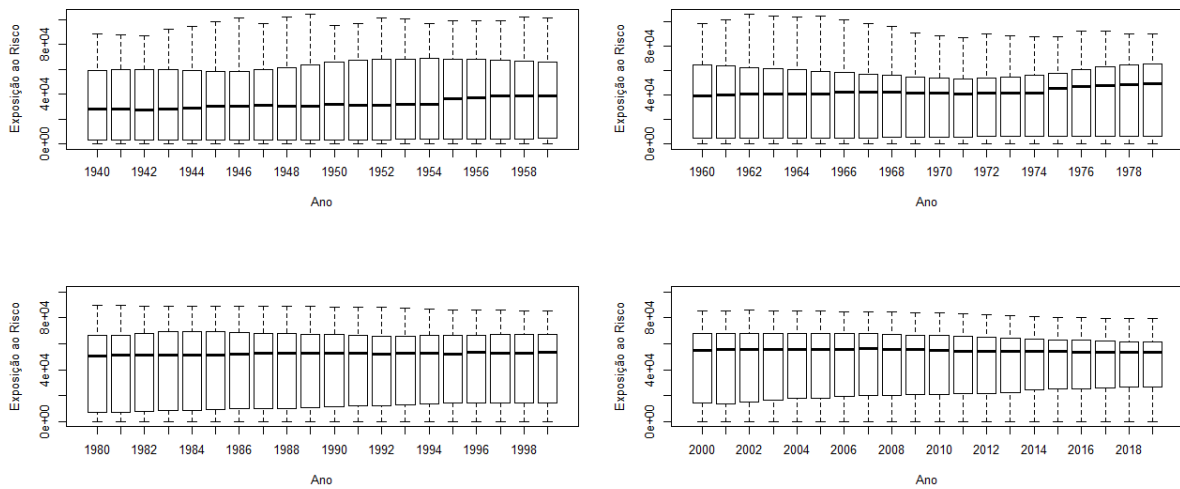


Figura 6.10: *Boxplots* do número de expostos ao risco nos anos 1940-2020 - Sexo Masculino.

6.1.2 Caso 2: Sexo Feminino

O número de mortes do sexo feminino de todos os anos do estudo pode ser observado a partir das Figuras A.9, A.10, A.11 e A.12 do Apêndice A. Tendo como exemplo os anos 1940, 1970, 2000 e 2020, através da Figura 6.11 observa-se que, tal como no caso anterior, o número de mortes nos primeiros anos (1940 e 1970) é mais elevado para idades entre 0-5 anos devido às razões apresentadas no Caso 1. A partir de 2000 a redução do número de mortes em idades iniciais foi mais acentuada, e os maiores valores passaram a ser verificados na faixa etária 80-90 anos. O maior valor observado foi em 2020 em mulheres com idade igual a 88 anos.

Na análise dos *boxplots* da Figura 6.12 constata-se que estes dados apresentam mais *outliers* do que os dados do sexo masculino, no entanto, tal como os *outliers* dos dados do Caso 1, depois de analisados um a um, estes também não são considerados erros mas sim valores extremos, sendo também considerados no estudo. Neste caso a faixa etária onde se encontram os *outliers* é mais ampla do que no Caso 1, sendo esta dos 73 aos 95 anos. Isto pode ser explicado através da esperança média de vida, uma vez que as mulheres apresentam uma maior esperança de vida do que os homens, logo têm tendência a morrer com mais idade.

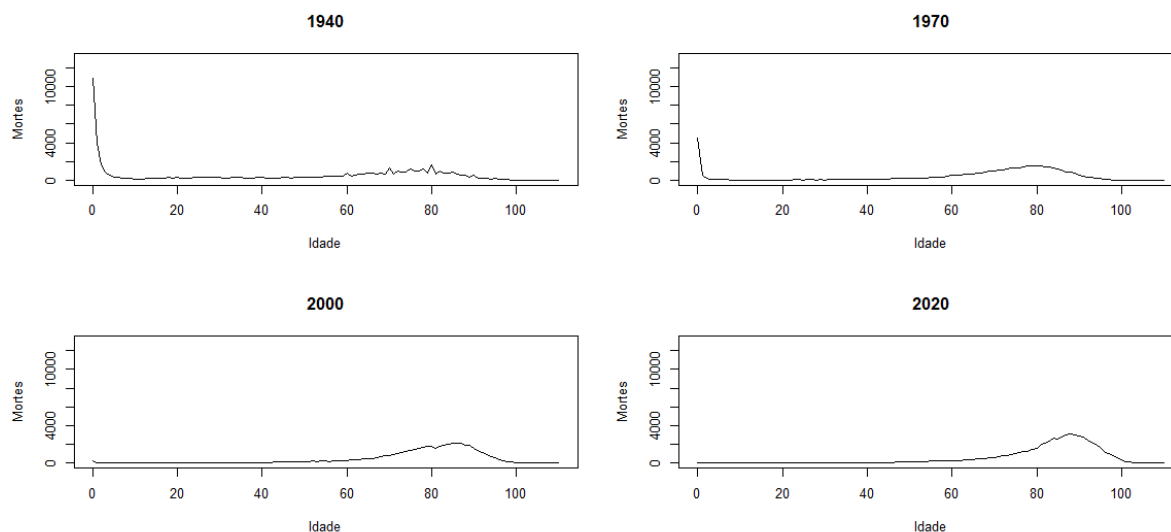


Figura 6.11: Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sexo Feminino.

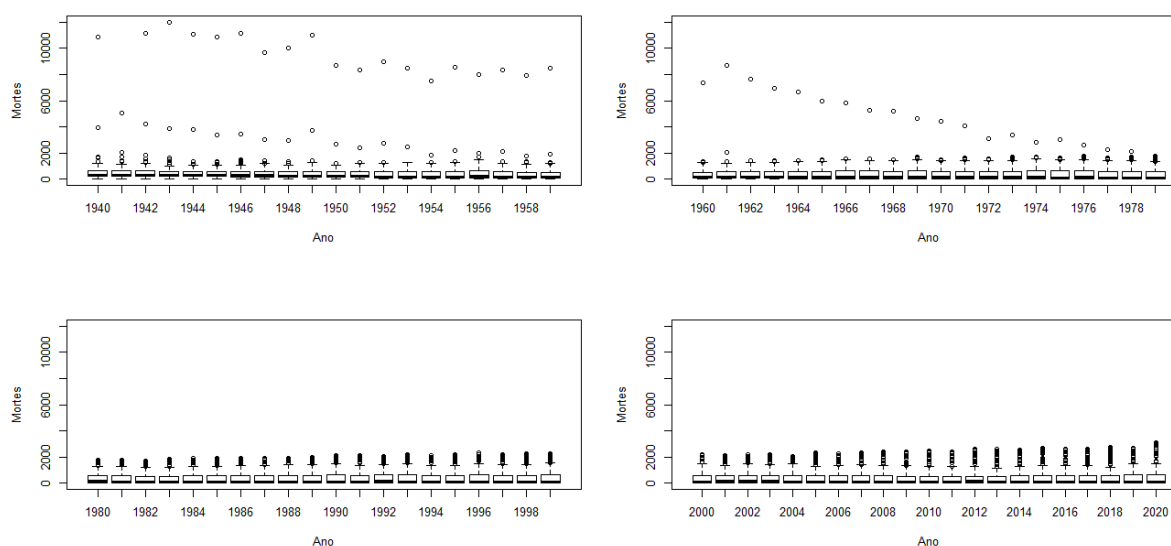


Figura 6.12: *Boxplots* do número de mortes nos anos 1940-2020 - Sexo Feminino.

Na Figura 6.13 é apresentado um *zoom* dos *boxplots* referentes ao número de mortes, de forma a analisar a dispersão dos dados ao longo dos anos. Verifica-se que, em todos os anos do estudo, os dados apresentam uma assimetria positiva, uma vez que a cauda direita é mais alongada do que a cauda esquerda e, observa-se que a partir de 1960 a dispersão dos dados aumentou, tornando-se quase constante ao longo dos restantes anos do estudo. Relativamente à mediana, constata-se que o número de mortes diminuiu nos primeiros 20 anos e nos 60 anos seguintes este valor é muito semelhante.

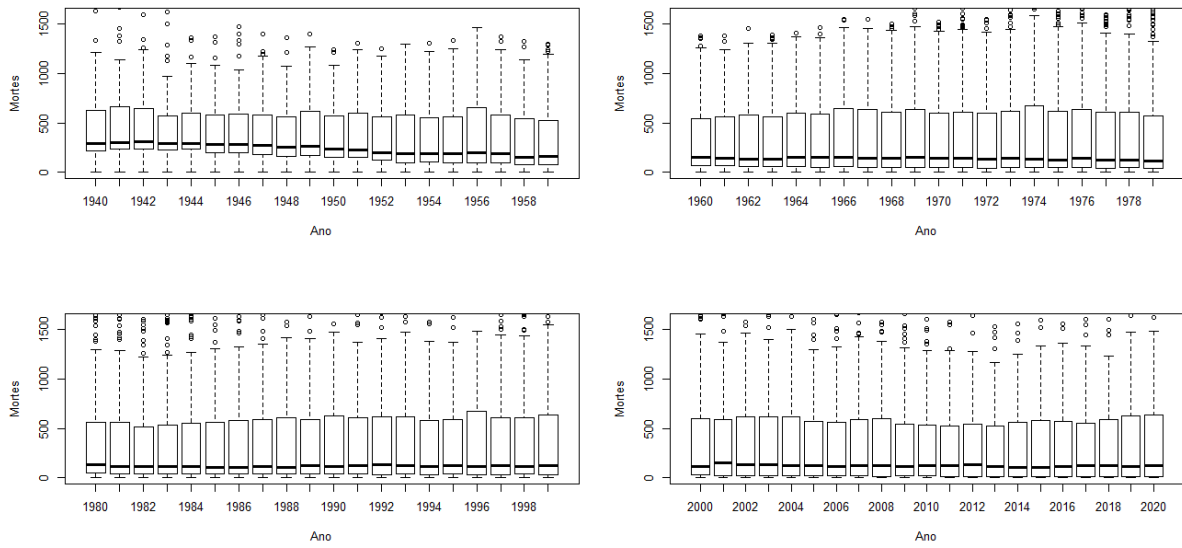


Figura 6.13: Zoom dos *boxplots* do número de mortes nos anos 1940-2020 - Sexo Feminino.

A respeito da exposição ao risco, nas Figuras A.13, A.14, A.15 e A.16 do Apêndice A estão apresentados os gráficos de todos os anos do estudo. Na Figura 6.14, a título de exemplo, observa-se o comportamento desta variável nos anos 1940, 1970, 2000 e 2020 e deteta-se que em 1940 e 1970 o maior valor de exposição ao risco era referente à idade 5 e 6 anos, respetivamente, e esse valor decrescia até aos 90 anos, sendo depois quase constante. Porém, nos últimos anos do estudo, estes valores sofreram algumas alterações, de modo que os maiores valores observados foram em mulheres com 25 anos em 2000 e 44 anos em 2020.

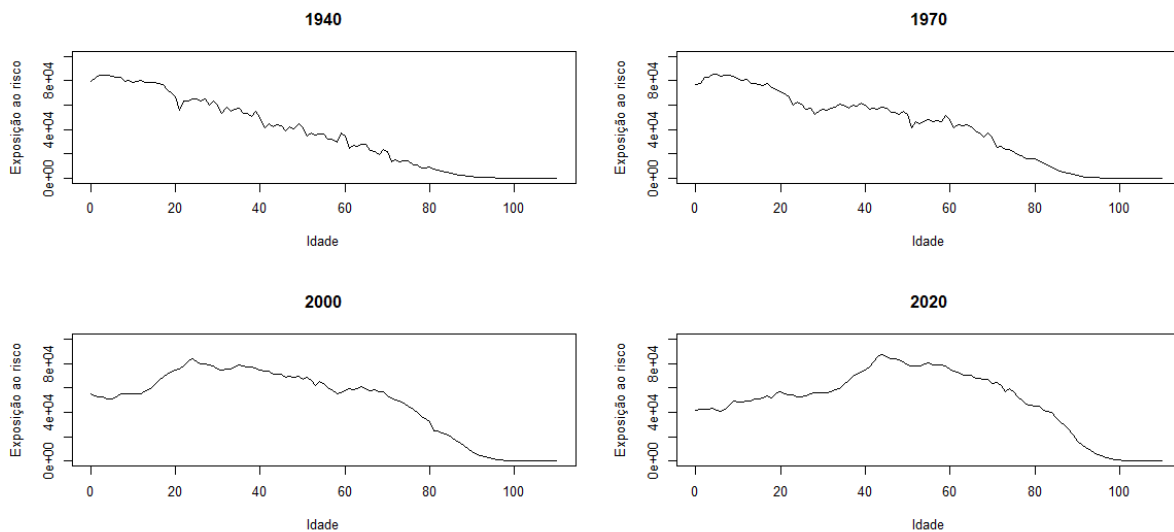


Figura 6.14: Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sexo Feminino.

Quanto à dispersão dos dados, na Figura 6.15 verifica-se que o ano que apresenta maior dispersão é o ano 1962 e, tal como no sexo masculino, o maior valor ocorreu na idade 0 deste mesmo ano. Analisando os *boxplots* desta figura percebe-se que os dados de todos os anos apresentam assimetria negativa, com a exceção dos anos 2015-2020 que apresentam assimetria positiva, o que significa que, exceto nos anos 2015-2020, as caudas esquerdas são mais alongadas do que as caudas direitas. Quanto à dispersão dos dados observa-se que esta diminui ao longo dos anos e, através da mediana, constata-se que o número de expostos ao risco apresenta algumas oscilações nos primeiros anos, porém nos últimos 20 anos este número tem sido praticamente constante.

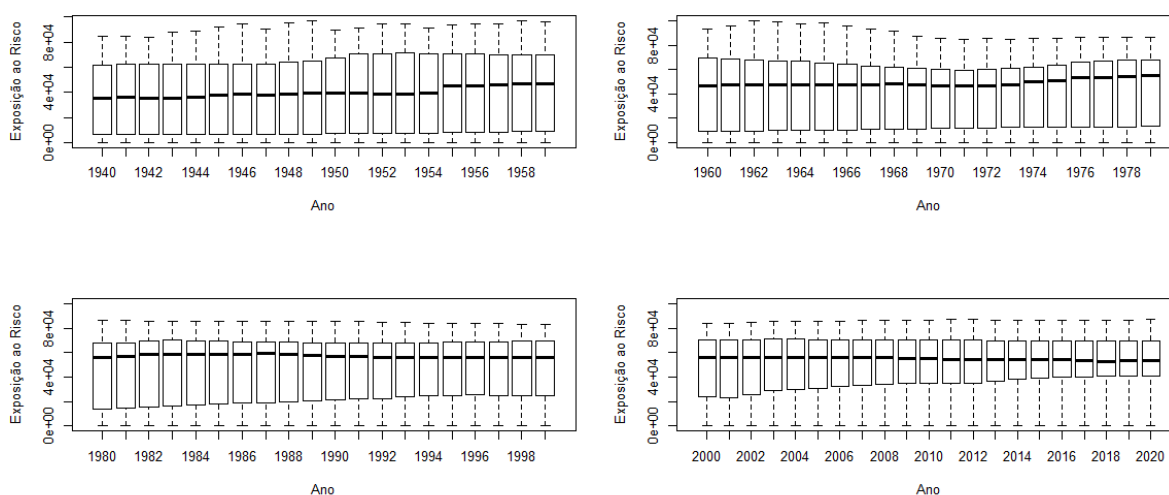


Figura 6.15: *Boxplots* do número de expostos ao risco nos anos 1940-2020 - Sexo Feminino.

6.1.3 Caso 3: Sem Distinção de Sexo

Nesta subsecção vão ser apresentados os dados sem distinção de sexo, ou seja, a junção dos dois casos anteriores. Nas Figuras A.17, A.18, A.19 e A.20 do Apêndice A estão disponíveis os gráficos referentes ao número de mortes ao longo de todos os anos do estudo. A título de exemplo, na Figura 6.16, estão apresentados os números de mortes ao longo das idades nos anos 1940, 1970, 2000 e 2020 e constata-se que esta variável apresenta valor mais elevados nos primeiros anos do estudo (1940 e 1970), como já seria expectável depois de ser realizada a análise dos dois primeiros casos, e, nos últimos anos, em 2000 e 2020, a mortalidade em idades entre 75-91 anos passa a ser a mais elevada, tendo atingido o máximo em 2020 em pessoas com 87 anos.

Relativamente à existência de *outliers*, na Figura 6.17 verifica-se que os dados apresentam valores

discrepantes porém, como foi referido anteriormente, estes são considerados como valores extremos. Estes valores correspondem a idades iniciais desde 1940 até 1979 e, entre 1980-2020 estão compreendidos na faixa etária 75-91 anos.

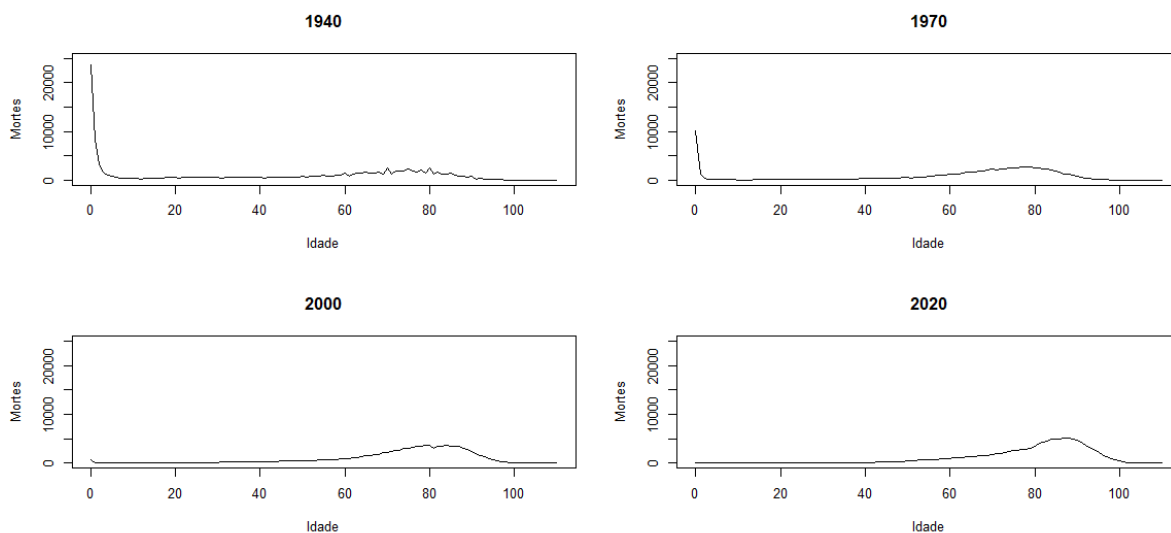


Figura 6.16: Número de mortes nos anos 1940, 1970, 2000 e 2020 - Sem Distinção de Sexo.

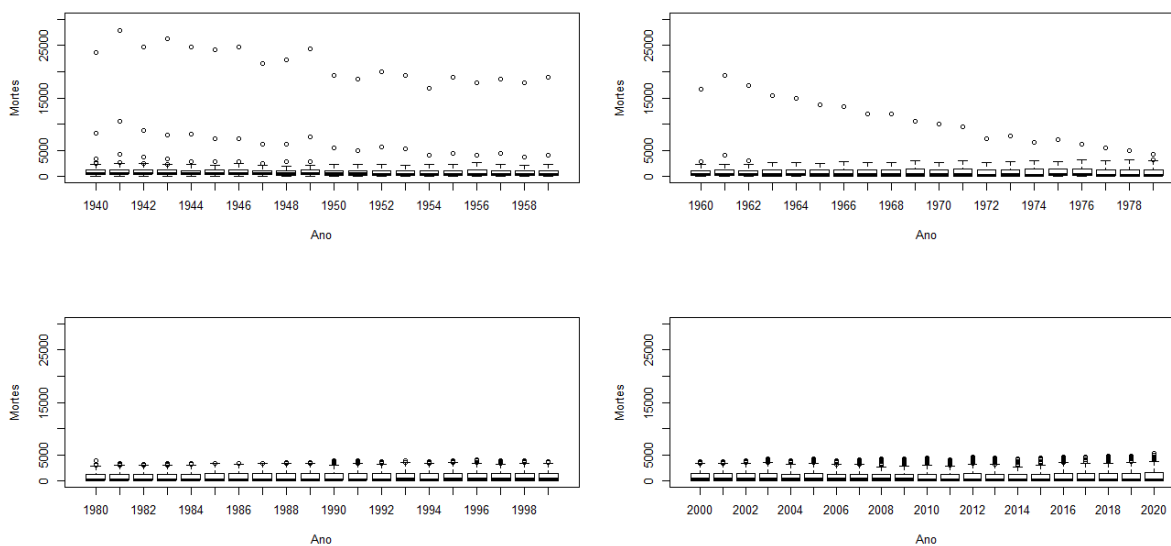


Figura 6.17: *Boxplots* do número de mortes nos anos 1940-2020 - Sem Distinção de Sexo.

De modo a analisar a dispersão dos dados, na Figura 6.18 é apresentado o *zoom* dos *boxplots* do número de mortes e a partir desta constata-se que a dispersão dos dados aumentou ao longo dos anos, porém através da mediana entende-se que o número de mortes diminuiu nos primeiros 20 anos e nos 60

anos seguintes tem-se mantido praticamente constante. Relativamente à assimetria dos dados, observa-se em todos os anos uma assimetria positiva, isto é, a cauda direita é mais alongada do que a cauda esquerda.

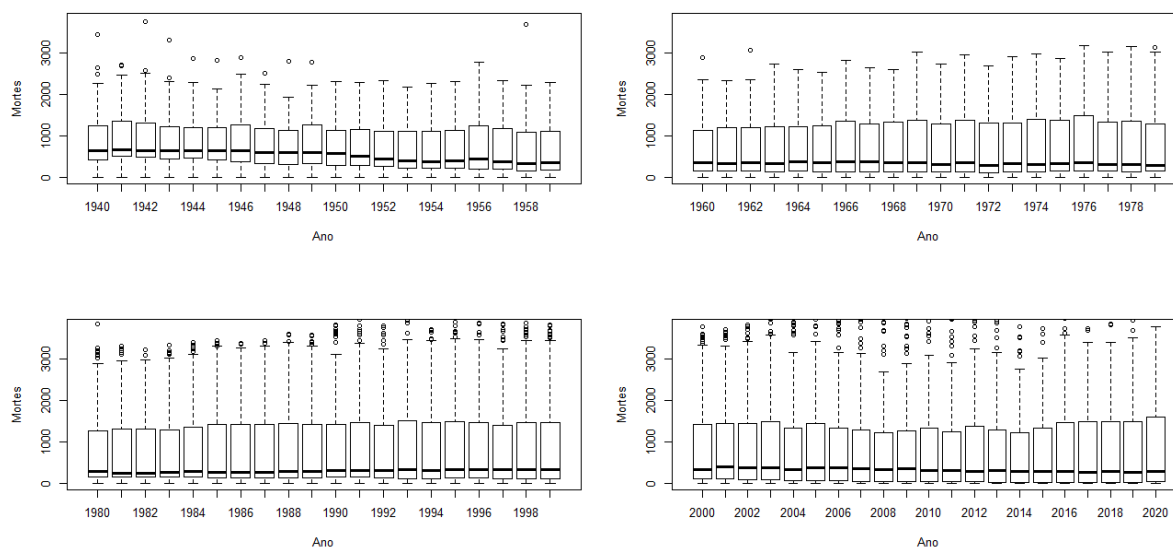


Figura 6.18: Zoom dos *boxplots* do número de mortes nos anos 1940-2020 - Sem Distinção de Sexo.

O comportamento do número de expostos ao risco ao longo dos anos do estudo pode ser observado nas Figuras A.21, A.22, A.23 e A.24 do Apêndice A. Na Figura 6.19 estão apresentados, como exemplo, os gráficos dos anos 1940, 1970, 2000 e 2020 e a partir desta é possível constatar que em 1940 e 1970 este valor decresce de 0 a 90 anos, sendo que a partir dos 90 anos é quase constante. Em 2000 e 2020, o comportamento desta variável altera-se ao longo dos anos, atingindo como valor máximo 24 e 44 anos, respetivamente.

Para estudar a distribuição dos dados foram elaborados os *boxplots* representados na Figura 6.20 onde se observa que, à exceção dos anos 1950-1954, os dados de todos os anos apresentam uma assimetria negativa, ou seja, a cauda esquerda é mais alongada do que a cauda direita, e os dados que apresentam uma maior amplitude são referentes ao ano de 1962. Tal como nos casos anteriores, os dados não apresentam nenhum *outlier*. Constata-se também que a dispersão dos dados tem diminuído ao longo dos anos e que a exposição ao risco tem sido praticamente constante nos últimos 40 anos.

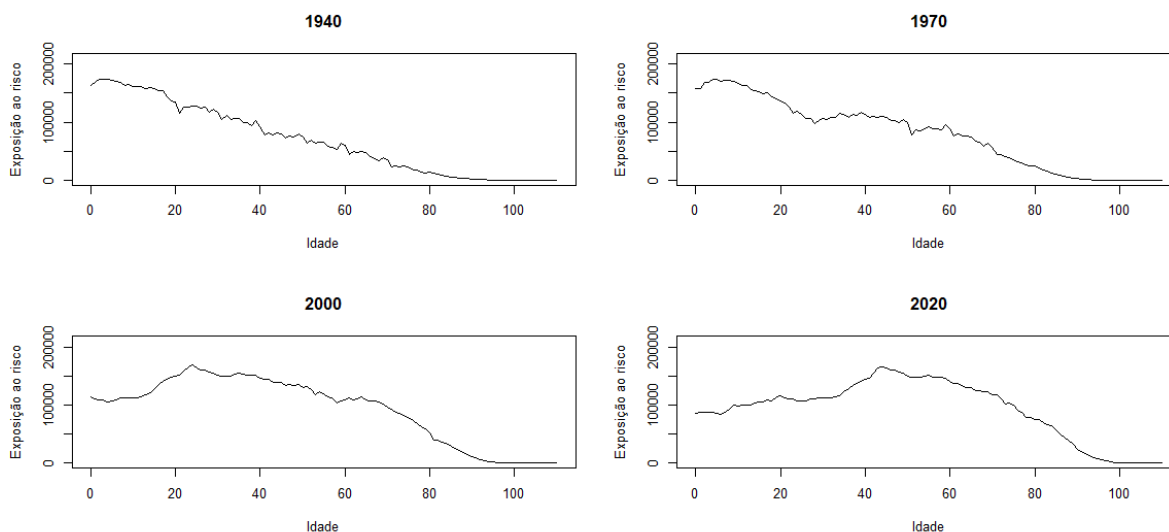


Figura 6.19: Número de expostos ao risco nos anos 1940, 1970, 2000 e 2020 - Sem Distinção de Sexo.

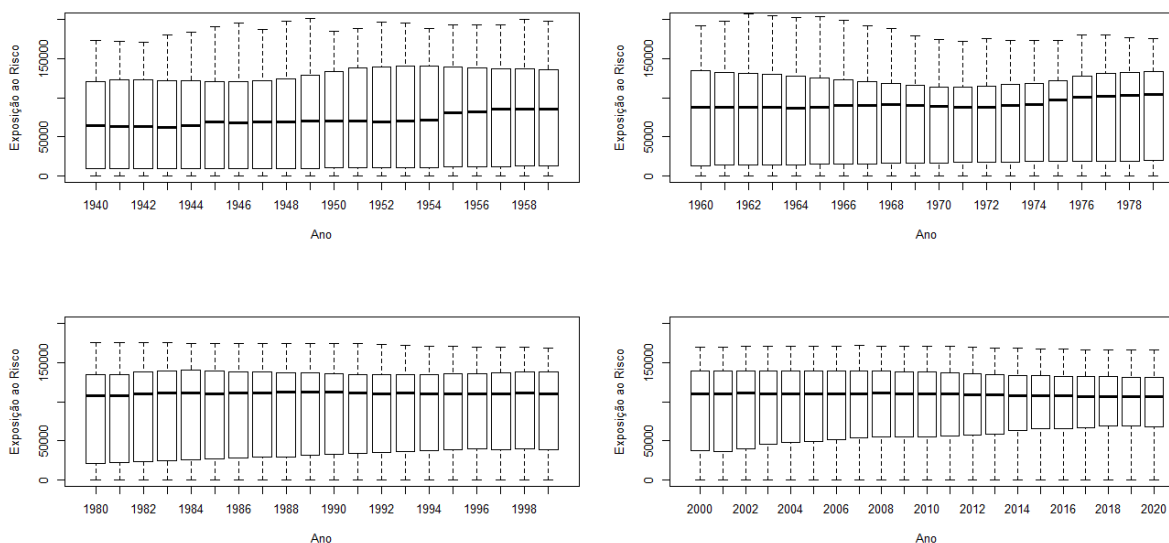


Figura 6.20: *Boxplots* do número de expostos ao risco nos anos 1940-2020 - Sem Distinção de Sexo.

De forma a complementar as análises gráficas apresentadas anteriormente, foram calculadas as medidas descritivas de cada variável para cada sexo e estas estão apresentadas nas Tabelas A.1 e A.2 do Apêndice A. Para verificar a evolução de cada variável ao longo do tempo, foram também calculados os valores médios das mesmas em cada ano, para cada sexo (Tabelas A.3 e A.4 do Apêndice A). Uma vez que os valores obtidos sem distinção de sexo são a junção do sexo masculino e feminino, neste caso são apenas comparados os dois últimos.

Comparando os resultados do sexo masculino e feminino em relação à variável de exposição ao risco, constata-se que apesar do valor mais elevado ter sido atingido no sexo masculino, em média, o sexo feminino apresenta valores superiores desta variável em comparação com o sexo masculino. Ao analisar as tabelas referentes às médias de cada ano, observa-se tanto no sexo masculino como no feminino, em geral, uma tendência crescente ao longo do domínio temporal, porém, nos anos 1970 e 2020 deteta-se um decréscimo desta variável em ambos os sexos. De forma a analisar-se a dispersão dos dados, foi calculado o desvio padrão e verifica-se que os dados do sexo masculino apresentam uma maior variabilidade comparativamente com os do sexo feminino.

Relativamente ao número de mortes, observa-se que, apesar do valor mínimo observado ser igual em ambos os sexos, o sexo masculino apresenta o valor mais elevado desta variável e que, em média, apresenta valores superiores em cada ano comparativamente com o sexo feminino, com a exceção do ano 2020. No que se refere à tendência observada desta variável, no caso do sexo masculino, verifica-se uma tendência decrescente entre 1940 e 1970 e a partir de 1980 verifica-se uma tendência crescente até 2020, existindo apenas uma quebra em 2010. No caso do sexo feminino observa-se um comportamento semelhante, a única diferença é que a tendência decrescente observa-se de 1940 a 1980 e a partir desse ano a tendência passa a ser crescente até 2020, não existindo nenhuma quebra durante esses anos. Por fim, no que diz respeito à dispersão dos dados, observa-se que os dados do sexo masculino apresentam uma maior dispersão em comparação com os do sexo feminino.

6.2 Aplicação do Método Lee-Carter

Como foi mencionado no Capítulo 3, o Método Lee-Carter (LC) é aplicado aos dados para calcular as previsões das taxas de mortalidade e é definido por

$$\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}.$$

Este método foi utilizado nos dados dos três casos abordados na Secção 6.1. Inicialmente foi necessário ajustar o modelo demográfico, ou seja, estimar os parâmetros da equação do modelo 3.1, α_x , β_x e k_t , sendo que α_x indica o nível médio das taxas de mortalidade no tempo para cada idade x , β_x representa as oscilações nas taxas de mortalidade na idade x , em resposta a alterações no nível geral de mortalidade k_t e k_t constitui o índice de nível geral da mortalidade no ano t . Neste caso os parâmetros foram estimados pelo método dos mínimos quadrados com a decomposição em valores

singulares (SVD). Nas Tabelas B.1, B.3, B.5 do Apêndice B encontram-se as estimativas dos parâmetros α_x e β_x nos casos com e sem distinção de sexo. Uma vez que o parâmetro k_t necessitou de ser reestimado, devido às estimativas iniciais serem calculadas a partir da matriz de logaritmos das taxas de mortalidade e não pela matriz das taxas de mortalidade, apenas é relevante apresentar os resultados dessa reestimação, Tabelas B.2, B.4 e B.6 do Apêndice B. Nas Figuras 6.21, 6.22 e 6.23 observa-se o comportamento das estimativas de cada parâmetro ao longo das idades (de 0 a 100 anos), no caso de α_x e β_x , e ao longo do tempo (de 1940 a 2020), no caso de k_t .

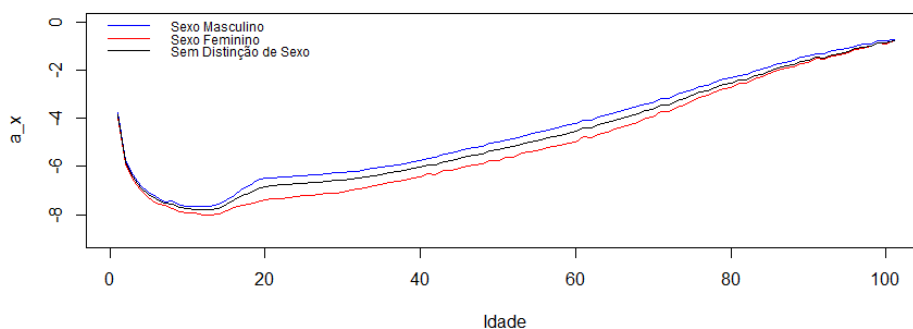


Figura 6.21: Estimativas do nível médio das taxas de mortalidade, α_x , ao longo das idades 0-100 anos.

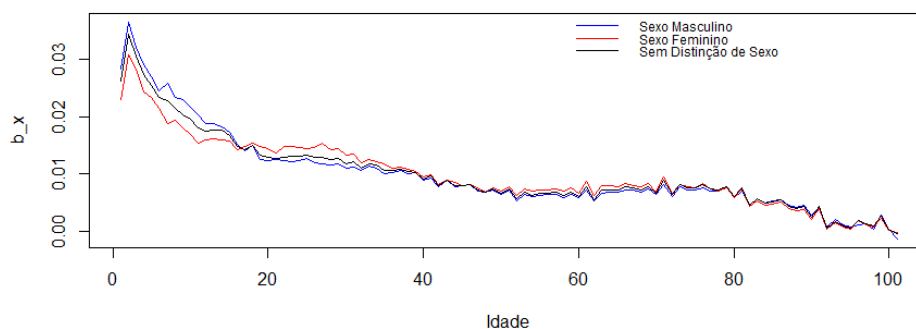


Figura 6.22: Estimativas das oscilações nas taxas de mortalidade, β_x , ao longo das idades 0-100 anos.

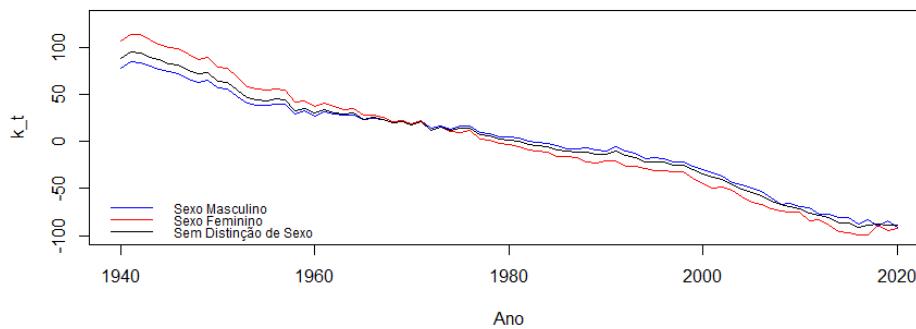


Figura 6.23: Estimativas do nível geral de mortalidade, k_t , nos anos 1940-2020.

A curva dos $\hat{\alpha}_x$ apresenta um comportamento que já se esperava, isto é, exibe valores mais elevados de taxa de mortalidade para recém-nascidos, de seguida estes valores decrescem atingindo um valor mínimo nas idades 10 e 11 anos no caso dos Homens, 12 e 13 no caso das Mulheres e Sem Distinção de Sexo e, por fim, é notório um crescimento até à idade 100 anos, sendo este o valor máximo deste parâmetro nos 3 casos. Relativamente aos gráficos das estimativas dos parâmetros $\hat{\beta}_x$ e \hat{k}_t , inicialmente verificam-se valores elevados de $\hat{\beta}_x$ em idades jovens que revelam que quanto menor for a idade, maior será a sensibilidade à variação de \hat{k}_t e, ao longo das restantes idades este valor decresce. Outra conclusão que se pode retirar sobre os valores de $\hat{\beta}_x$ é que se estes valores forem baixos significa que as taxas de mortalidade variam pouco quando o nível geral de mortalidade se altera. Por fim, os valores de \hat{k}_t representam o nível geral de mortalidade e estes apresentam uma tendência negativa ao longo dos anos, com a exceção do ano 2020 que, devido ao aparecimento da doença por coronavírus (COVID-19), apresentou uma tendência crescente, uma vez que uma consequência desta doença é muitas vezes a morte. Por esta razão decidiu-se não utilizar os dados do ano 2020 para projetar as tábuas de mortalidade.

6.3 Modelação do Nível Geral de Mortalidade - Modelação ARIMA

O objetivo deste estudo não incluía a previsão de dados porém, no caso da empresa pretender realizar previsões da taxa de mortalidade, decidiu-se realizar a modelação do nível geral de mortalidade de forma a obter o melhor modelo de previsão para o caso do sexo masculino, feminino e sem distinção de sexo.

A evolução do nível geral de mortalidade k_t no tempo é uma série temporal e é necessário modelar

o comportamento da mesma e, posteriormente, realizar projeções para ser possível projetar a taxa de mortalidade para anos futuros. Nesta secção utiliza-se a metodologia Box-Jenkins proposta por Lee-Carter, contudo, apesar dos autores recomendarem a utilização de um modelo ARIMA(0,1,0), foram testados outros modelos para perceber qual seria o mais adequado. As séries temporais do índice k_t foram analisadas em três casos referentes ao tipo de sexo, isto é, sexo masculino, sexo feminino e sem distinção de sexo. Os dados originais foram divididos em dois conjuntos, um de treino (ou estimação) e um de teste (ou previsão), com o intuito de avaliar a precisão dos diferentes modelos de previsão aplicados. O período de treino inclui as observações entre os anos 1940 e 2010, e o período de teste compreende as observações de 2011 a 2019.

6.3.1 Identificação do Modelo

Inicialmente, é necessário observar a série temporal da reestimação do parâmetro k_t . Observando as Figuras 6.24, 6.25 e 6.26 constata-se que as séries apresentam uma variância quase constante porém, não são constantes na média. Logo, a partir desta análise pode-se suspeitar que as séries são não estacionárias. Outra conclusão que pode ser retirada após a observação das mesmas figuras é que nenhuma das séries apresenta sazonalidade, isto é, não apresenta nenhum padrão que se repete na série temporal em períodos específicos, e por essa razão será selecionado um modelo ARIMA e não SARIMA.

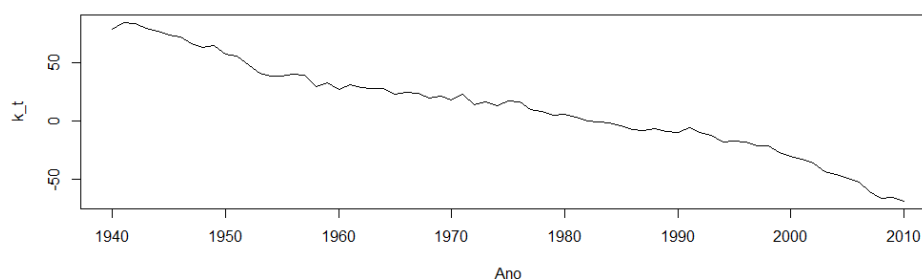


Figura 6.24: Série temporal do parâmetro k_t no período de treino (1940-2010) - Sexo Masculino.

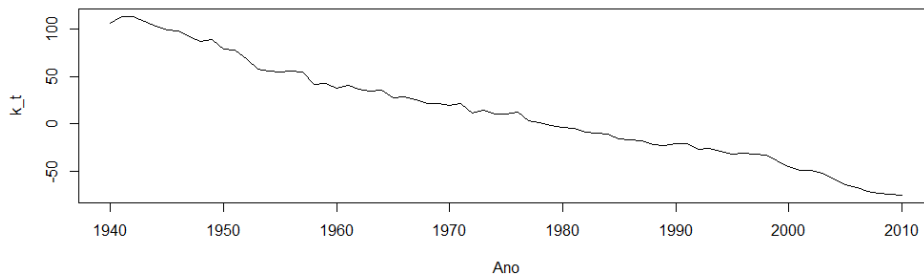


Figura 6.25: Série temporal do parâmetro k_t no período de treino (1940-2010) - Sexo Feminino.

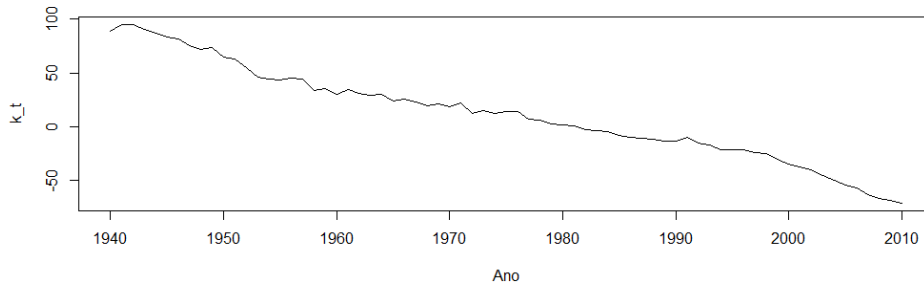


Figura 6.26: Série temporal do parâmetro k_t no período de treino (1940-2010) - Sem Distinção de Sexo.

Outra forma de averiguar a estacionariedade da série temporal passa por analisar a FAC empírica e utilizar os testes ADF e KPSS. Nas Figuras 6.27, 6.28 e 6.29 é possível observar que as FAC apresentam um decaimento muito lento para zero à medida que k (lag) aumenta, por isso confirma-se novamente que as séries são não estacionárias.

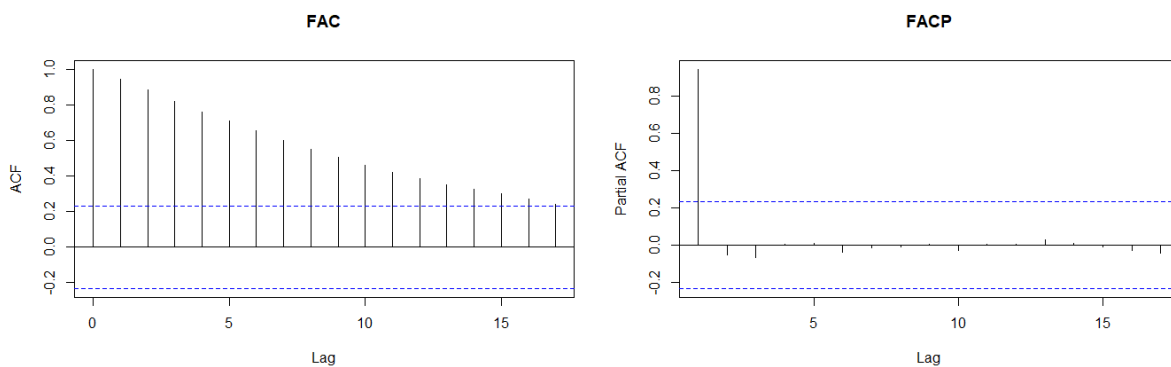


Figura 6.27: Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Masculino.

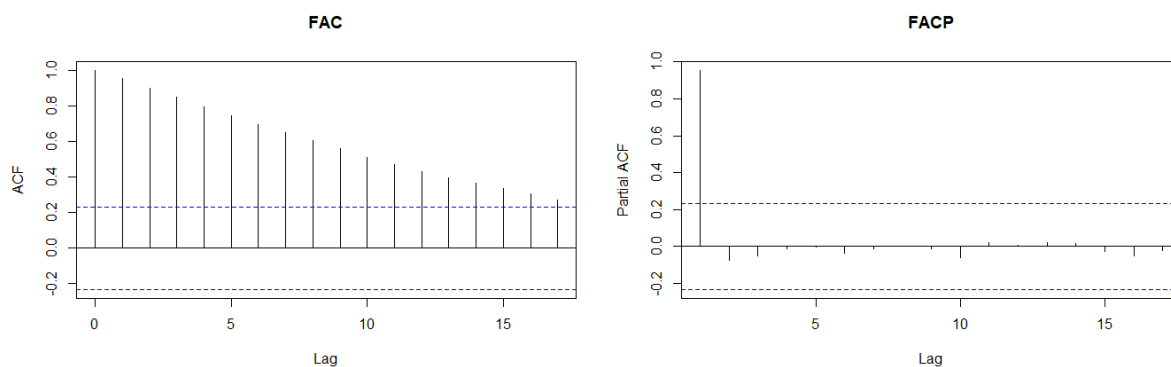


Figura 6.28: Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Feminino.

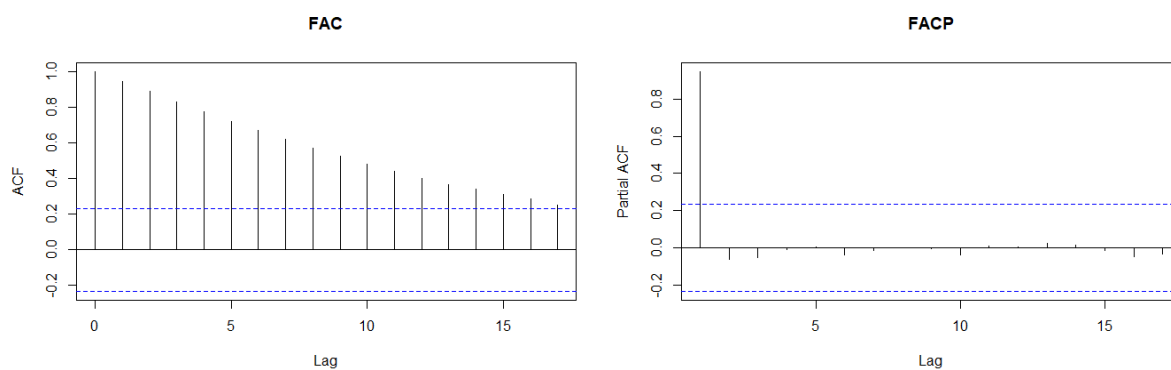


Figura 6.29: Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sem Distinção de Sexo.

Relativamente aos testes ADF e KPSS, para analisar se a série é estacionária, a estatística de teste terá de ser inferior ao valor crítico que, para um nível de significância de 5% é -1,95 no teste ADF e 0,463 no teste KPSS. De realçar que caso o teste ADF indique que a série é não estacionária e o teste KPSS indique que a série é estacionária, considera-se que a série é estacionária em tendência, ou seja, a tendência precisa de ser removida para que a série seja rigorosamente estacionária. No caso do sexo masculino a estatística de teste é igual a 0,7896 para o teste ADF e 0,7146 para o teste KPSS, no sexo feminino as estatísticas de teste são 0,1558 no teste ADF e 0,1342 no teste KPSS e, no caso de sem distinção de sexo as estatísticas de teste são 0,6642 para o teste ADF e 0,1024 para o teste KPSS. Assim, pode-se concluir que, com um nível de significância de 5%, as séries são não estacionárias.

Uma vez que todas as séries são não estacionárias, é necessário realizar diferenciações. Primeiramente será realizada uma primeira diferenciação e, de seguida, será testada novamente a

estacionariedade das séries através dos métodos utilizados anteriormente.

Nas Figuras 6.30, 6.31 e 6.32 encontram-se as séries após a 1ª diferenciação e as respectivas FAC e FACP empíricas. Após a observação destes gráficos parece que as séries já são estacionárias, porém é ainda necessário confirmar através dos testes ADF e KPSS. No caso do teste ADF as estatísticas de teste são -1,1335, -1,3025 e -1,2348 para o sexo masculino, feminino e sem distinção de sexo, respetivamente. Relativamente ao teste KPSS, as estatísticas de teste são 0,1699, 0,0848 e 0,1148 no caso do sexo masculino, feminino e sem distinção de sexo, respetivamente. Tendo em atenção o que foi referido anteriormente sobre o caso de os dois testes levarem a diferentes conclusões, podemos concluir que, mesmo depois de ser realizada uma diferenciação às séries e com um nível de significância de 5%, estas continuam a ser não estacionárias. Por esta razão é necessário realizar mais uma diferenciação ($d=2$).

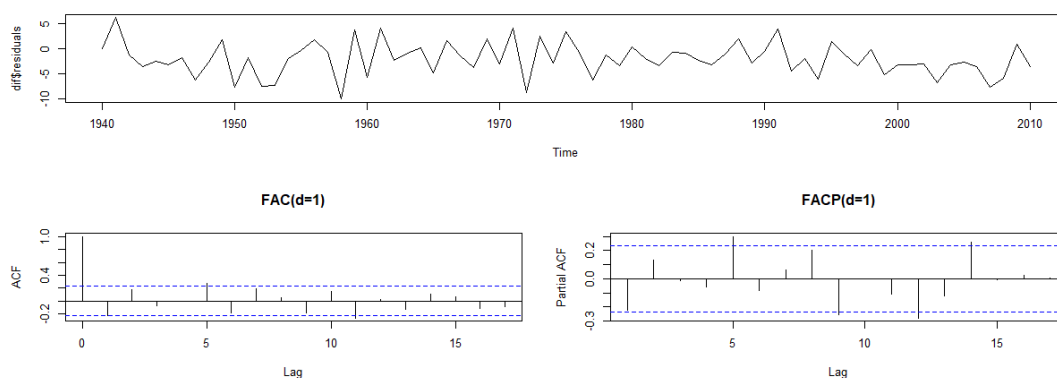


Figura 6.30: Série do nível geral de mortalidade k_t após a 1ª diferenciação ($d=1$) e respetivas FAC e FACP estimadas - Sexo Masculino.

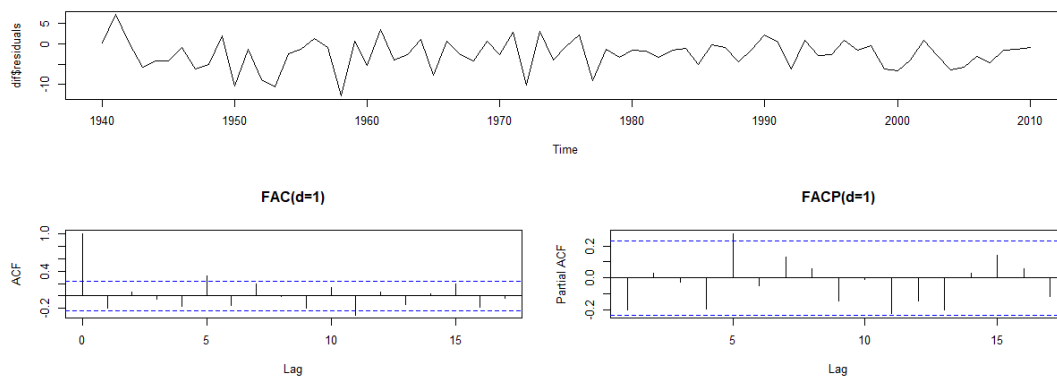


Figura 6.31: Série do nível geral de mortalidade k_t após a 1ª diferenciação ($d=1$) e respetivas FAC e FACP estimadas - Sexo Feminino.

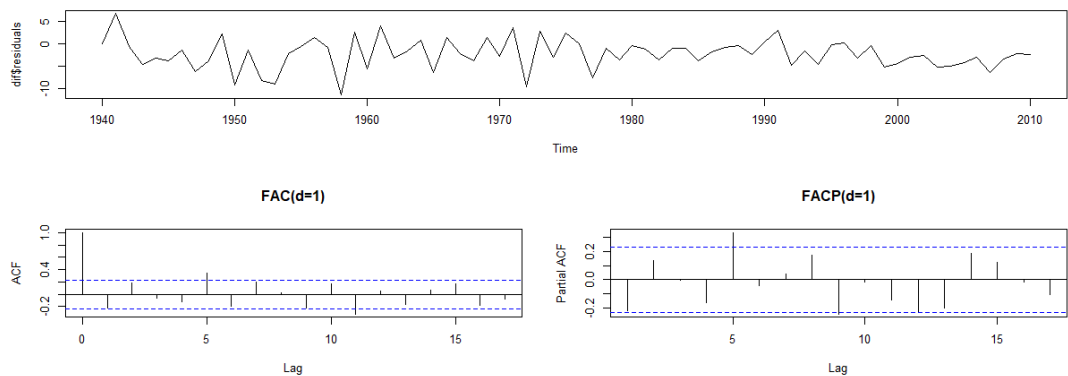


Figura 6.32: Série do nível geral de mortalidade k_t após a 1ª diferenciação ($d=1$) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.

Os gráficos das séries após a 2ª diferenciação e as respectivas FAC e FACP empíricas estão apresentadas nas Figuras 6.33, 6.34 e 6.35. Mais uma vez, a partir da análise gráfica, parece que as séries são estacionárias, contudo é fundamental realizar os testes para confirmar. Depois de realizados os testes ADF e KPSS foram obtidas as seguintes estatísticas de teste: -2,5201 (no sexo masculino), -2,5356 (no sexo feminino) e -2,5953 (sem distinção de sexo) no caso do teste ADF e 0,0471 (no sexo masculino), 0,2029 (no sexo feminino) e 0,1740 (sem distinção de sexo) no caso do teste KPSS. Sendo assim, uma vez que todas as estatísticas de teste são inferiores aos respectivos valores críticos, pode-se concluir que, com um nível de significância de 5%, as séries são estacionárias após a 2ª diferenciação.

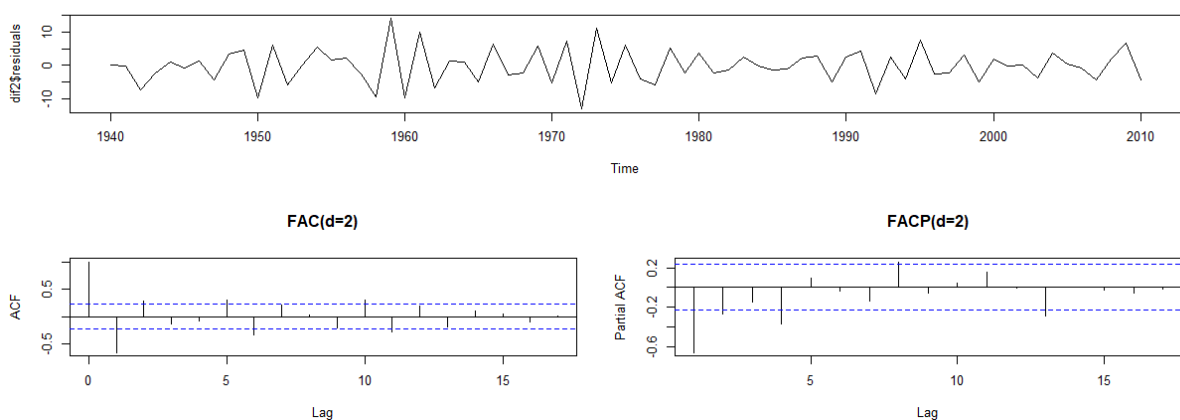


Figura 6.33: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respectivas FAC e FACP estimadas - Sexo Masculino.

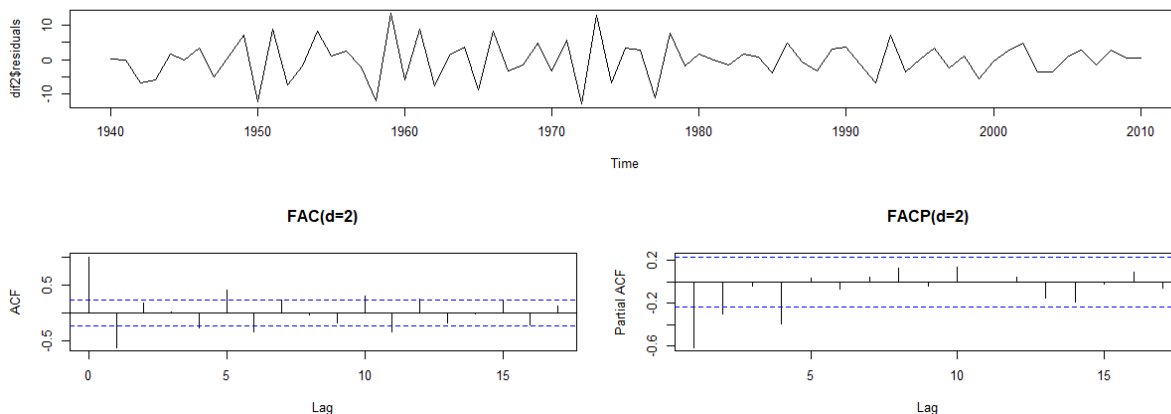


Figura 6.34: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respectivas FAC e FACP estimadas - Sexo Feminino.

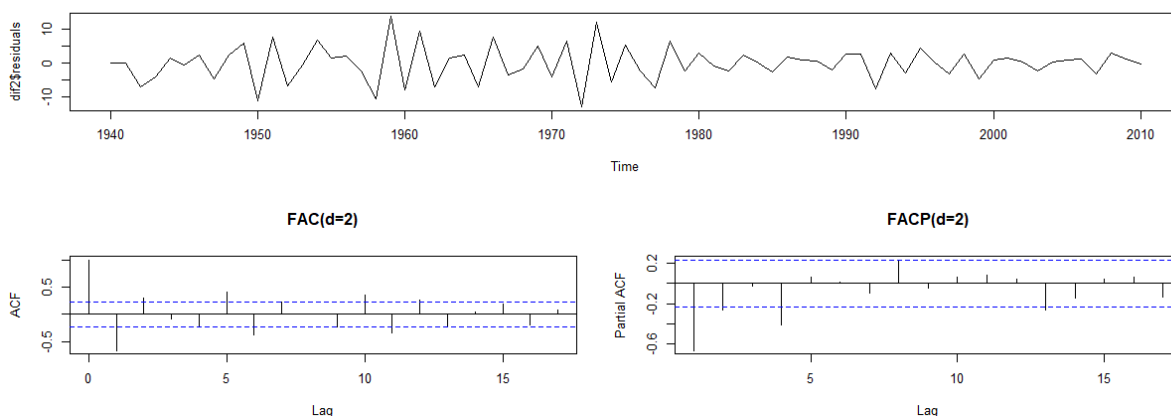


Figura 6.35: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.

Após concluída a primeira sub-etapa da identificação do modelo, falta apenas escolher as ordens p e q de modo a ajustar um modelo ARIMA. Para a escolha das ordens p e q do modelo, foi utilizado o algoritmo Hyndman-Khandakar (Hyndman & Athanasopoulos, 2021) que consiste em considerar quatro modelos iniciais, ARIMA(0,d,0), ARIMA(2,d,2), ARIMA(1,d,0) e ARIMA(0,d,1), e, destes quatro modelos, aquele que apresentar o menor AIC e BIC é considerado como o "modelo atual", que será posteriormente comparado com variações desse mesmo modelo que apresentam variações de ± 1 nas ordens p e q . O modelo escolhido será aquele que apresentar um menor AIC e BIC, sendo que poderá ser o "modelo atual" ou uma das suas variações.

Nas Tabelas 6.4, 6.5 e 6.6 estão apresentados os modelos “concorrentes” de cada fase do algoritmo e os respectivos parâmetros, sendo que a ordem d é sempre igual a 2 uma vez que foram necessárias duas diferenciações em cada série para as tornar estacionárias. Na 1ª fase do algoritmo os modelos que apresentam um menor valor de AIC e BIC são ARIMA(0,2,1) tanto no sexo masculino, como no sexo feminino e no caso sem distinção de sexo. Depois de realizadas as variações deste modelo em cada um dos casos (2ª fase do algoritmo), os modelos mais adequados são ARIMA(1,2,1) para o sexo masculino e sem distinção de sexo e, no caso do sexo feminino, o modelo ARIMA(0,2,1).

Tabela 6.4: Identificação das ordens do modelo ARIMA - Sexo Masculino.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	425,88	428,12
	ARIMA(2,2,2)	-0,45*	0,00*	-0,75*	-0,13*	371,54	382,71
	ARIMA(1,2,0)	-0,68	-	-	-	387,06	391,53
	ARIMA(0,2,1)	-	-	-1,00	-	371,54	376,01
2ª Fase	ARIMA(0,2,2)	-	-	-1,23	0,36	367,99	374,69
	ARIMA(1,2,1)	-0,34	-	-0,88	-	367,65	374,36
	ARIMA(1,2,2)	0,32*	-	-1,56	0,76	369,18	378,11

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Tabela 6.5: Identificação das ordens do modelo ARIMA - Sexo Feminino.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	437,66	439,89
	ARIMA(2,2,2)	-1,03	-0,27	-0,23*	-0,77	385,88	397,05
	ARIMA(1,2,0)	-0,62	-	-	-	406,66	411,13
	ARIMA(0,2,1)	-	-	-1,00	-	384,82	389,29
2ª Fase	ARIMA(0,2,2)	-	-	-1,24	0,24*	383,31	390,01
	ARIMA(1,2,1)	-0,24*	-	-1,00	-	383,13	389,83
	ARIMA(1,2,2)	-0,26*	-	-0,98*	-0,02*	385,13	394,06

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Tabela 6.6: Identificação das ordens do modelo ARIMA - Sem Distinção de Sexo.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	425,78	428,02
	ARIMA(2,2,2)	-0,10*	0,17*	-1,12*	0,12*	372,16	383,33
	ARIMA(1,2,0)	-0,67	-	-	-	386,70	391,17
	ARIMA(0,2,1)	-	-	-1,00	-	371,77	376,23
2ª Fase	ARIMA(0,2,2)	-	-	-1,20	0,31*	369,69	376,39
	ARIMA(1,2,1)	-0,33	-	-0,89	-	368,870	375,57
	ARIMA(1,2,2)	0,39	-	-1,62	0,83	370,42	379,36

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

6.3.2 Avaliação do Modelo

Para validar os modelos escolhidos na subsecção anterior só falta verificar a significância estatística dos parâmetros estimados e se os resíduos da estimação apresentam um comportamento análogo ao de um ruído branco Gaussiano.

Na Tabela 6.7 estão apresentados os parâmetros de cada modelo para cada sexo. Através da análise dos mesmos constata-se que todos os parâmetros são significativos, para um nível de significância de 5%

Tabela 6.7: Modelos ARIMA mais adequados para cada caso.

Sexo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
Masculino	ARIMA(1,2,1)	-0,34	-	-0,88	-	367,65	374,36
Feminino	ARIMA(0,2,1)	-	-	-1,00	-	384,82	389,29
Sem Distinção	ARIMA(1,2,1)	-0,33	-	-0,89	-	368,87	375,57

Para finalizar esta etapa falta apenas analisar o comportamento dos resíduos dos modelos, para verificar se estes apresentam um comportamento análogo ao de um ruído branco Gaussiano. Para isso são utilizados os testes de *Ljung-Box* para testar a independência dos resíduos e o teste *t* de *Student* para testar o valor médio.

No caso do sexo masculino, o histograma da Figura 6.36 sugere que os resíduos da série apresentam uma distribuição ligeiramente assimétrica à esquerda porém, mediante o teste *Lilliefors*, a hipótese nula

de normalidade não é rejeitada, tendo como valor de prova 0,648. Uma vez que a normalidade não foi rejeitada, é possível testar se os resíduos apresentam valor médio nulo a partir do teste t , resultando num valor de prova igual a 0,199. Além disso, através da representação gráfica da série dos resíduos, estes parecem oscilar uniformemente em torno de zero, o que leva a crer que a variância é constante. Sendo assim, conclui-se que não há evidência estatística para rejeitar a hipótese de que os resíduos apresentam uma distribuição Normal de média nula e variância constante. O último pressuposto a ser testado é o da não correlação através do teste de *Ljung-Box* que tem como hipótese nula a correlação nula dos resíduos, ou seja, a independência dos mesmos. Com o valor de prova obtido no teste, 0,122, com um nível de significância de 5% não se rejeita a hipótese nula, ou seja, verifica-se a independência dos erros. Posto isto, pode-se considerar que o modelo escolhido é adequado aos dados.

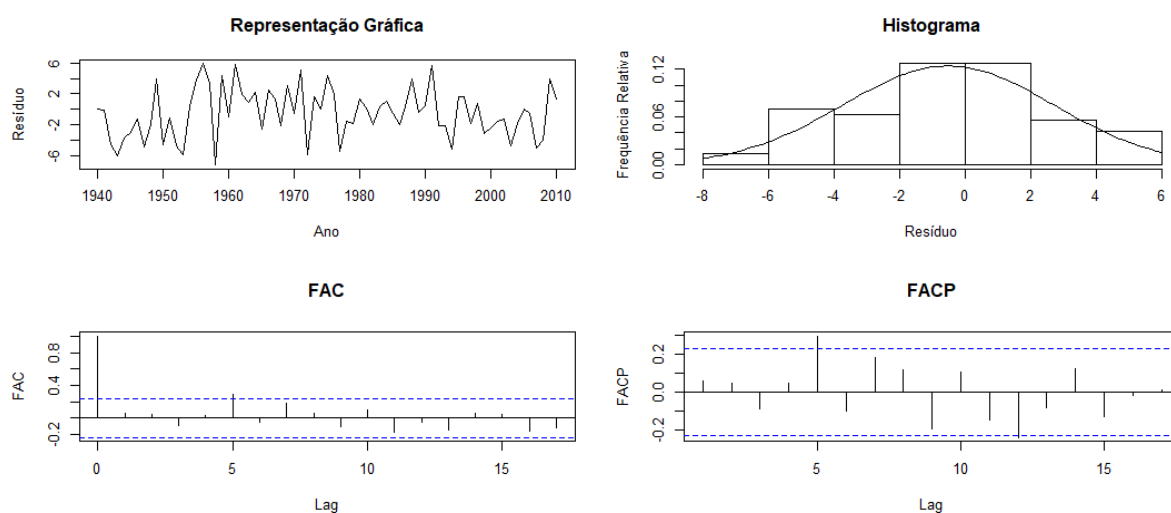


Figura 6.36: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.

Relativamente ao sexo feminino e sem distinção de sexo, estes casos não são tão simples quanto o primeiro. Nas Figuras 6.37 e 6.38 é possível verificar que em ambos os casos os resíduos apresentam uma distribuição ligeiramente assimétrica à esquerda, porém não é suficiente para que a partir do teste *Lilliefors* se rejeite a hipótese nula de normalidade, sendo que o valor de prova obtido neste teste são 0,150 no caso do sexo feminino e 0,865 no caso sem distinção de sexo. Contudo, no teste de *Ljung-Box*, a hipótese de correlação nula entre os resíduos é rejeitada com valores de prova 0,003 no sexo feminino e 0,021 no caso sem distinção de sexo, sendo por isso necessário recorrer ao modelo ARIMA alternativo em ambos os casos.

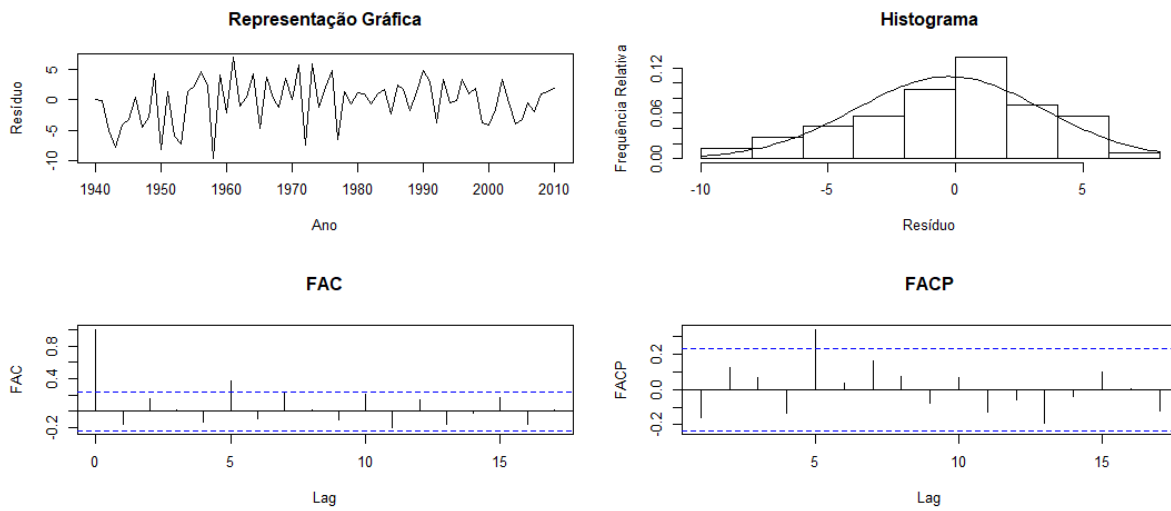


Figura 6.37: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.

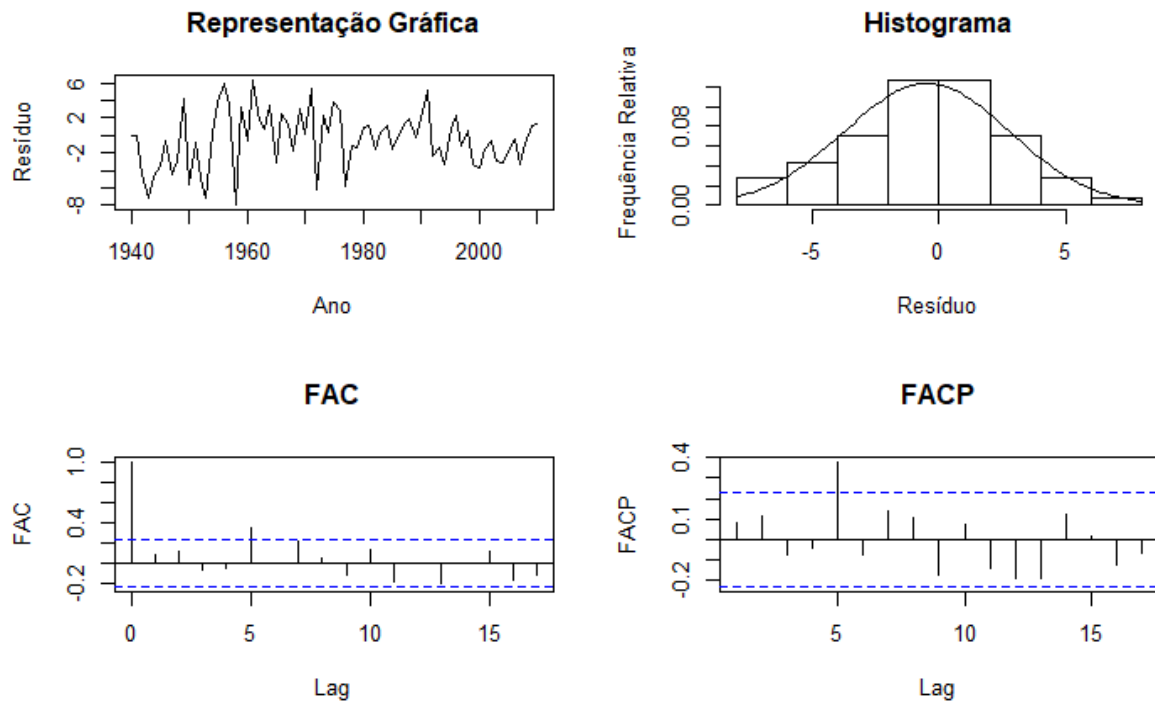


Figura 6.38: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.

A escolha do modelo alternativo consiste em escolher o segundo modelo com menor AIC e BIC e com parâmetros estatisticamente significativos. Para o sexo feminino, o modelo alternativo é o modelo

ARIMA(1,2,0) (Tabela 6.5) e para o caso sem distinção de sexo é o modelo ARIMA(0,2,1) (Tabela 6.6). Nas Figuras 6.39 e 6.40 apresenta-se a representação gráfica dos resíduos de cada modelo e os respectivos histograma, FAC e FACP. Através do histograma da Figura 6.39 constata-se que os resíduos apresentam uma distribuição assimétrica à esquerda mais acentuada do que no caso do histograma apresentado na Figura 6.40. Porém, utilizando o teste *Lilliefors*, prova-se que a hipótese de normalidade não é rejeitada em nenhum dos casos, tendo como valores de prova 0,068 no caso do sexo feminino e 0,864 no caso sem distinção de sexo. Em relação à existência de correlação entre os resíduos, através do teste *Ljung-Box*, verifica-se novamente que existe correlação em ambos os casos, apresentando como valor de prova 0,007 (sexo feminino) e 0,0001 (sem distinção de sexo). Apesar da normalidade e do valor médio nulo serem verificados em ambos os casos, uma vez que existe correlação entre os resíduos, nenhum dos modelos pode ser utilizado para modelação nem previsão.

Posto isto, foram testados os modelos seguidos com menor AIC e BIC. No caso do sexo feminino só foi possível testar o modelo ARIMA(0,2,0) uma vez que os restantes apresentavam coeficientes estatisticamente não significativos. Já no caso sem distinção de sexo foram testados os modelos ARIMA(1,2,2), ARIMA(1,2,0) e ARIMA(0,2,0). Em todos os modelos as hipóteses de normalidade e de valor médio nulo não foram rejeitadas, porém a hipótese de existência de correlação entre os resíduos foi sempre verificada.

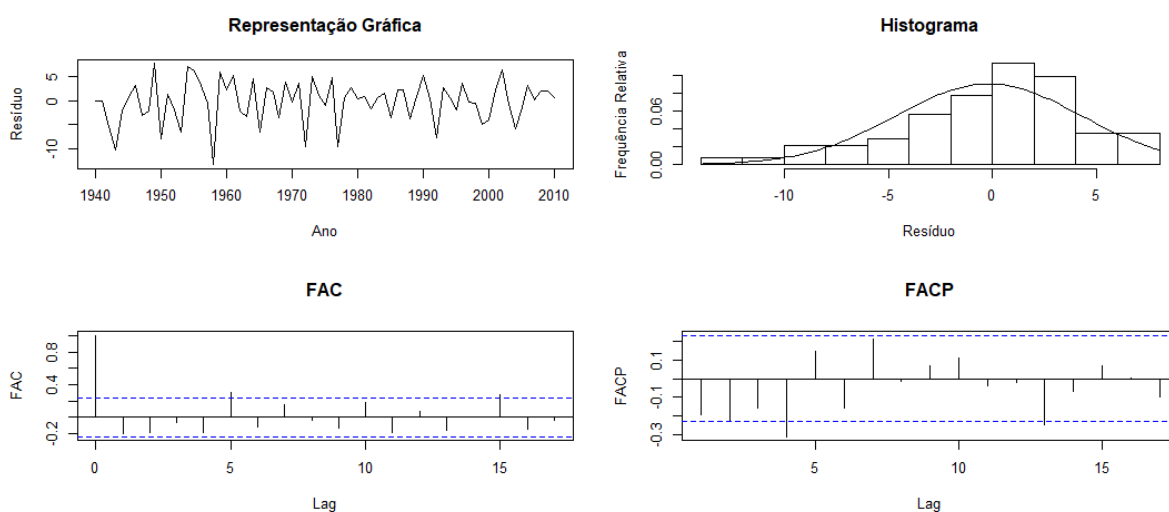


Figura 6.39: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sexo Feminino.

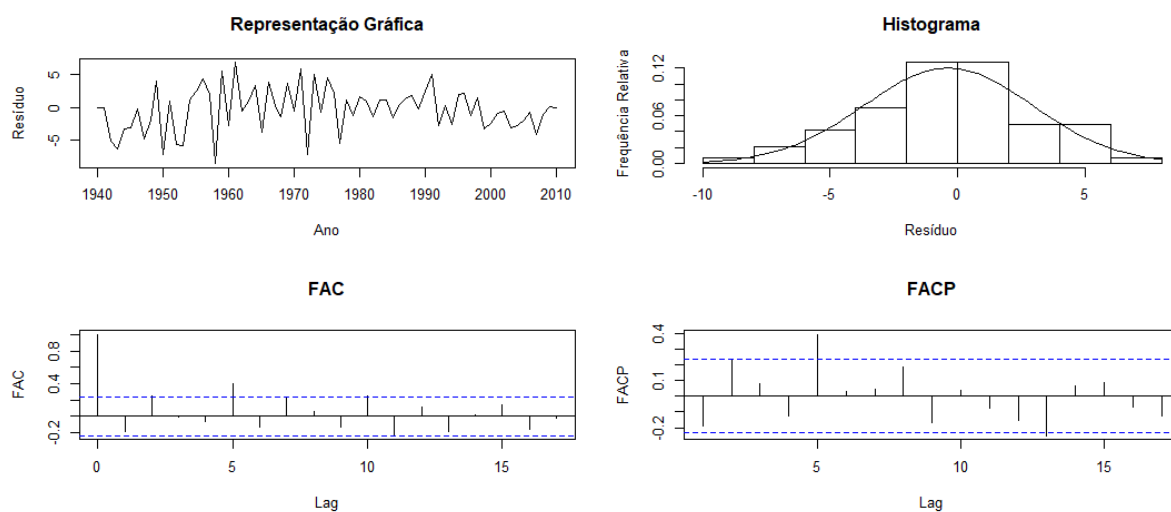


Figura 6.40: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sem Distinção de Sexo.

Uma vez que nenhum dos modelos obtidos seguindo o algoritmo Hyndman-Khandakar (Hyndman & Athanasopoulos, 2021) verifica os pressupostos necessários, é utilizada outra forma para identificar os parâmetros p e q do modelo, através das representações gráficas da FAC e FACP das Figuras 6.34 e 6.35. Através destas é possível reconhecer-se o comportamento de um modelo AR(4), ou, alternativamente, de um MA(1), tanto no sexo feminino como no caso sem distinção de sexo. No entanto, para além destes modelos, também foram ajustados modelos “vizinhos” para uma escolha mais cuidada.

Nas Tabelas 6.8 e 6.9 estão apresentados os modelos estudados para cada um dos casos e os respetivos valores de AIC e BIC. Mediante os resultados observa-se que, para ambos os sexos, os modelos que apresentam menores valores de AIC e BIC são os modelos ARIMA(2,2,3). Desta forma falta apenas averiguar se os resíduos de cada modelo verificam os pressupostos pretendidos.

Tabela 6.8: Identificação das ordens do modelo ARIMA - Sexo Feminino.

Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	AIC	BIC
ARIMA(1,2,3)	0,81	-	-	-	-2,08	1,34	-0,26*	386,88	398,05
ARIMA(2,2,3)	-1,63	-0,98	-	-	0,54	-0,60	-0,94	377,60	391,01
ARIMA(3,2,2)	0,04*	0,06*	-0,03*	-	-1,28*	0,28*	-	389,06	402,46
ARIMA(4,2,1)	-0,58	-0,39	-0,42	-0,44	-0,58	-	-	382,71	396,12

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Tabela 6.9: Identificação das ordens do modelo ARIMA - Sem Distinção de Sexo.

Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	AIC	BIC
ARIMA(1,2,3)	-0,40*	-	-	-	-0,78	-0,03*	-0,19*	371,52	382,69
ARIMA(2,2,3)	-1,63	-0,98	-	-	0,51	-0,60	-0,91	361,57	374,98
ARIMA(3,2,2)	0,37	-0,01*	-0,41	-	-1,67	1,00	-	364,40	377,81
ARIMA(4,2,1)	-0,69	-0,38*	-0,47	-0,47	-0,46	-	-	365,02	378,43

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Após a observação das Figuras 6.41 e 6.42 constata-se que os resíduos parecem apresentar uma distribuição ligeiramente assimétrica porém, esta não é significativa para que a hipótese de normalidade seja rejeitada através do teste *Lilliefors* (valor de prova igual a 0,369 no caso do sexo feminino e 0,862 no caso sem distinção de sexo), com um nível de significância de 5%. Relativamente à independência dos resíduos e ao valor médio dos mesmos, através do teste de *Ljung-Box* verifica-se que os resíduos são independentes no caso do sexo feminino, com um valor de prova igual a 0,170, e 0,231 para o sexo feminino e sem distinção de sexo, respetivamente, e, através do teste *t*, constata-se que os resíduos apresentam um valor médio nulo (valor de prova igual a 0,645 no sexo feminino e 0,379 no caso sem distinção de sexo). Por fim, a partir da observação do gráfico dos resíduos, estes parecem oscilar uniformemente em torno de zero, o que leva a crer que a variância é constante.

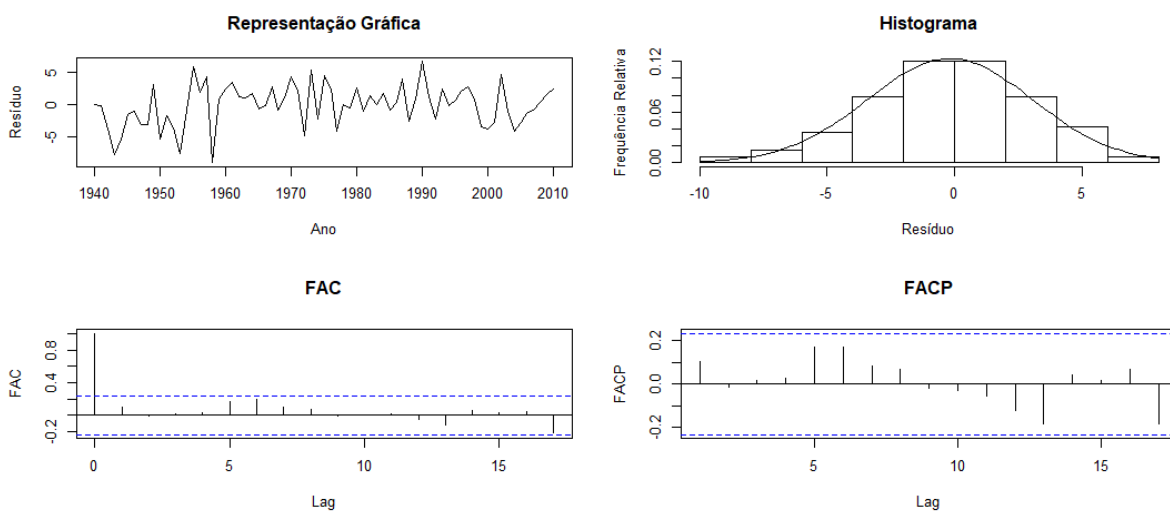


Figura 6.41: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.

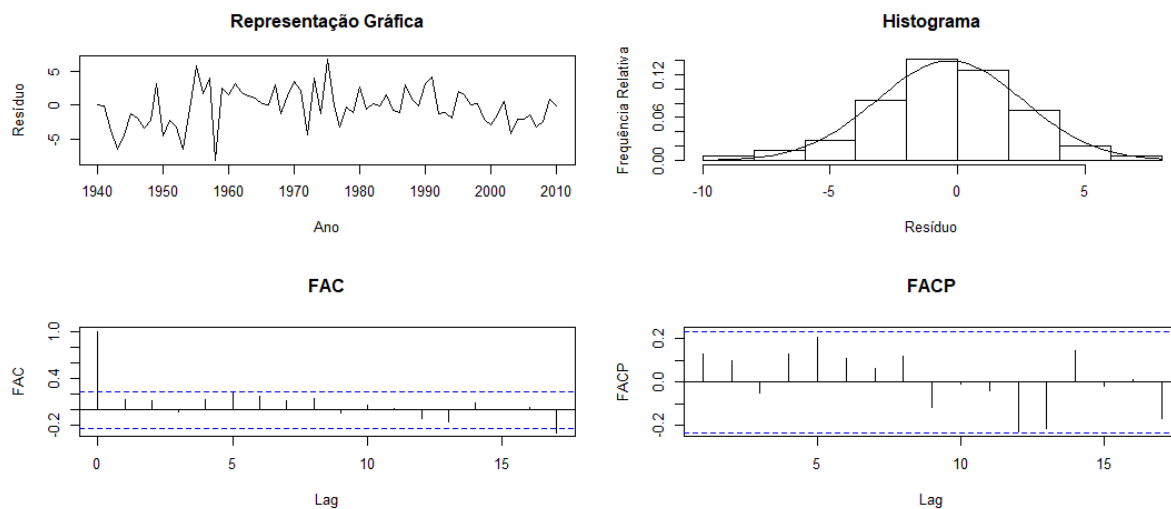


Figura 6.42: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.

Visto que foram verificados todos os pressupostos, é obter realizar as estimativas pontuais (no período de treino) e as previsões pontuais e intervalares (no período de teste) de cada caso, apresentadas nas Figuras 6.43, 6.44 e 6.45. Estas sugerem que a qualidade preditiva de cada modelo é melhor na série de treino do que na série de teste e isso pode ser comprovado através das medidas de avaliação calculadas para cada um dos casos e representadas na Tabela 6.10.

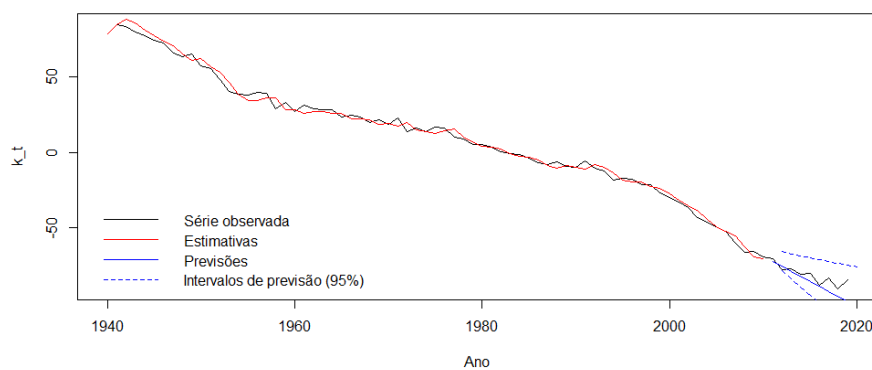


Figura 6.43: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.

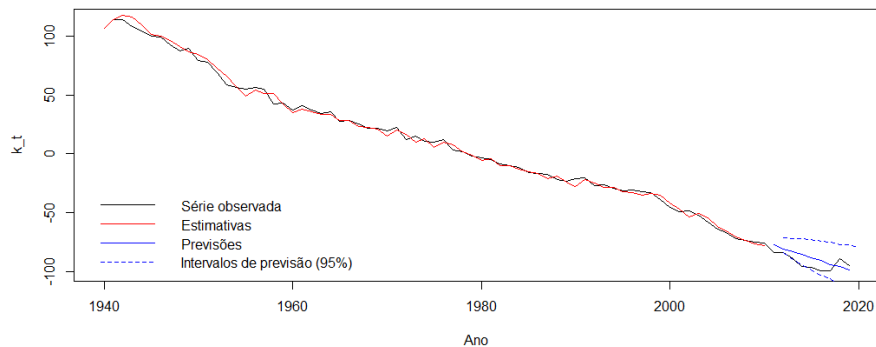


Figura 6.44: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.

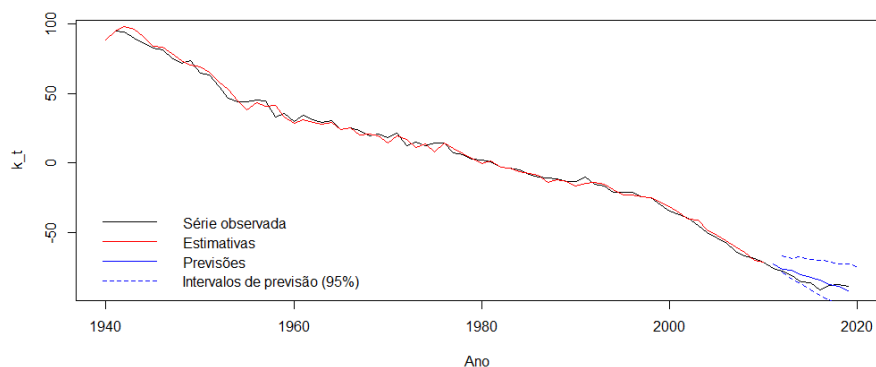


Figura 6.45: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.

Relativamente aos intervalos de previsão, no caso do sexo feminino 8 das 9 observações correspondentes à série de teste pertencem ao mesmo, sendo que a taxa de cobertura é igual a 88,89%, ou seja, inferior à esperada (95%) e, nos casos do sexo masculino e sem distinção de sexo, a taxa de cobertura de cada um é 100%, ou seja, todas as observações de cada série de teste pertencem ao respetivo intervalo de previsão.

Tabela 6.10: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.

Sexo	EQM	REQM	EPAM	EAM	U-Theil	
Masculino	10,427	3,229	0,315	2,617	0,600	
Feminino	10,310	3,211	0,116	2,529	0,742	Período de Treino
Sem Distinção	8,192	2,862	0,126	2,230	0,326	
Masculino	39,199	6,261	0,055	4,579	1,111	
Feminino	47,209	6,871	0,070	6,500	1,321	Período de Teste
Sem Distinção	14,457	3,802	0,039	3,333	1,266	

6.4 Modelação do Nível Geral de Mortalidade - Método de Holt

Tal como foi referido na Secção 5.8, o método de alisamento exponencial mais indicado para séries que não apresentam sazonalidade, o caso presente neste estudo, é o Método de Holt ou Método Linear de Holt. Na Tabela 6.11 estão apresentadas as estimativas iniciais para as constantes de alisamento e para o nível e declive, para cada sexo. Relativamente à constante de alisamento associada ao nível, $\hat{\alpha}$, observa-se que, para todos os sexos, este valor é elevado, o que indica que é dado mais peso às observações mais recentes do que às mais distantes e que o nível muda rapidamente com a finalidade de capturar a série de alta tendência. Já os valores da constante de alisamento associada ao declive, $\hat{\beta}$, dos três casos são bastante próximos de 0, isto sugere que a tendência quase não muda com o tempo, ou seja, que o declive num dado instante fica determinado, quase na totalidade, pelo valor no instante anterior.

Tabela 6.11: Estimativas iniciais para o nível e declive e estimativas das constantes de alisamento, resultantes da aplicação do Método de Holt à série temporal do nível geral de mortalidade k_t produzido a partir do Método Lee-Carter.

Sexo	Estimativas			
	$\hat{\alpha}$	$\hat{\beta}$	\hat{l}_1	\hat{b}_1
Masculino	0,642	0,114	83,832	-1,829
Feminino	0,786	0,0001	117,167	-2,773
Sem Distinção	0,722	0,072	93,828	-2,368

Nas Figuras 6.46, 6.47 e 6.48 encontram-se as representações gráficas das estimativas pontuais no período de treino e as previsões pontuais e intervalares do período de teste. Observando apenas as figuras, parece que as estimativas são melhores do que as previsões e isso pode ser comprovado através do cálculo das medidas de avaliação para ambos os períodos.

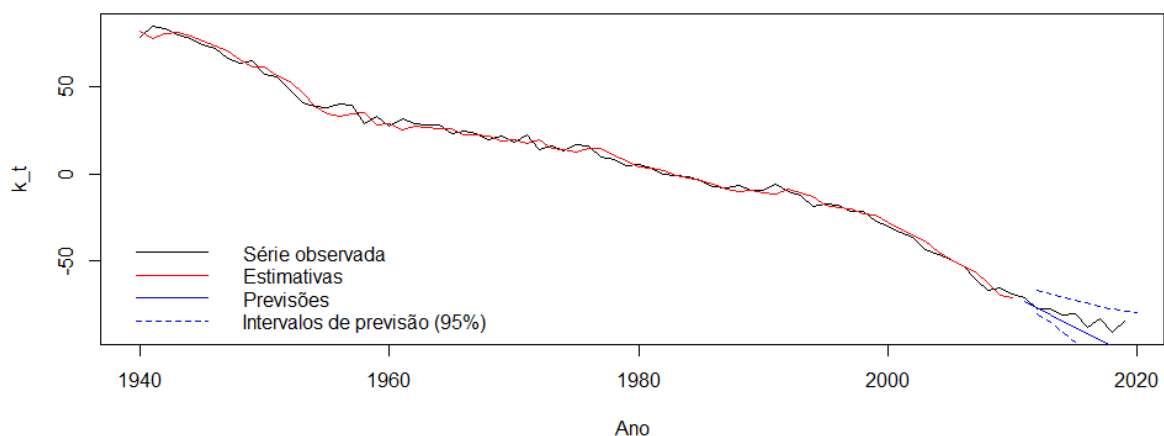


Figura 6.46: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.

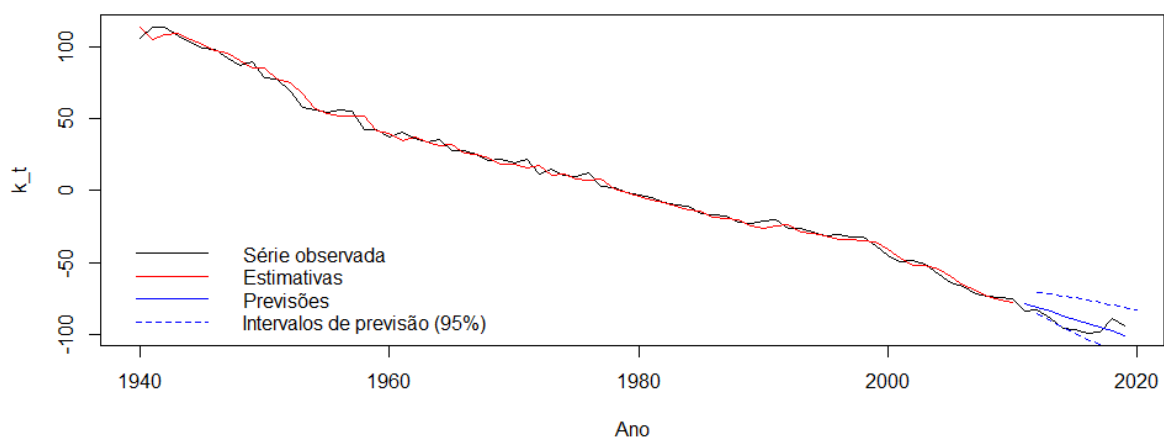


Figura 6.47: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.

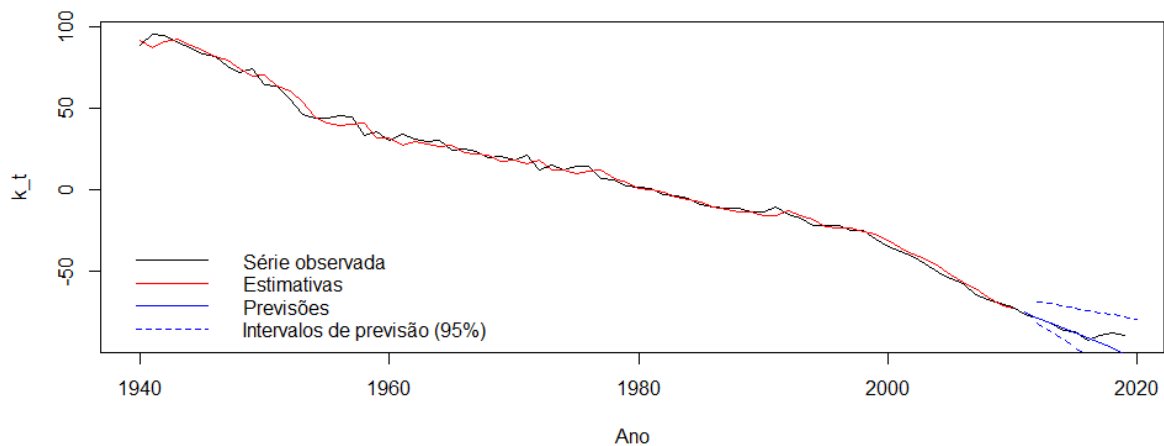


Figura 6.48: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.

Na Tabela 6.12 estão apresentadas as medidas de avaliação calculadas e, com base nas mesmas, é possível comprovar a premissa anterior, uma vez que as medidas de avaliação têm valores menores na série de treino do que na série de teste. Apesar das medidas de avaliação apresentarem valores elevados no período de teste, todas as observações da série de teste pertencem aos intervalos de previsão nos três casos de estudo, tendo por isso uma taxa de cobertura de 100%.

Tabela 6.12: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.

Sexo	EQM	REQM	EPAM	EAM	U-Theil	
Masculino	10,487	3,238	0,293	2,636	0,607	Período de Treino
Feminino	13,090	3,618	0,122	2,854	0,517	
Sem Distinção	10,732	3,276	0,144	2,611	0,484	
Masculino	73,374	8,507	0,081	6,686	1,520	Período de Teste
Feminino	39,187	6,260	0,063	5,858	1,214	
Sem Distinção	25,883	5,088	0,038	3,372	1,676	

6.5 Extensão do Método Lee-Carter ou Modelo de Poisson

Como foi referido no Capítulo 3, a extensão do Método Lee-Carter foi proposta devido aos autores do mesmo acreditarem que o ajustamento dos parâmetros através do método dos mínimos quadrados via SVD não era o mais correto. Posto isto, a única diferença entre o Método Lee-Carter e a sua extensão é que no último método considera-se que o número de óbitos segue uma distribuição de Poisson e os parâmetros do mesmo são estimados utilizando o método de máxima verosimilhança. Nesta secção são apresentados todos os resultados deste método e é feita a comparação entre a extensão e o Método Lee-Carter.

Tal como o Método Lee-Carter, este método também foi aplicado aos dados descritos na Secção 6.1, com e sem distinção de sexo. A primeira etapa é o ajustamento do modelo demográfico, realizado através do método de máxima verosimilhança. Nas Figuras 6.49, 6.50, 6.51, 6.52, 6.53 e 6.54 estão representadas as estimativas dos parâmetros da equação do modelo 3.1, α_x , β_x e k_t , obtidas através do método dos mínimos quadrados (Método LC) e através do método de máxima verosimilhança (Extensão do Método LC). Após a observação das mesmas, constata-se que o parâmetro α_x é praticamente igual em todos os sexos, utilizando ambos os métodos. Já os parâmetros β_x e k_t , apesar de apresentarem um comportamento semelhante ao longo do tempo, exibem valores bastante diferentes, sendo que estes valores são mais dispares quando se trata do parâmetro k_t .

Posto isto pode-se afirmar que, tal como na secção 6.2, a curva dos $\hat{\alpha}_x$ apresenta um comportamento expectável, apresentando em ambos os métodos valores mínimos nas idades 10 e 11 anos no caso do sexo masculino e 12 e 13 anos nos casos do sexo feminino e sem distinção de sexo e valores máximos na idade 100 anos. Em relação ao parâmetro $\hat{\beta}_x$, utilizando a Extensão do Método LC, em idades iniciais o valor deste parâmetro é mais elevado, mais no sexo masculino e no caso sem distinção de sexo, significando que nestes casos a sensibilidade à variação do parâmetro \hat{k}_t é ainda mais elevada do que aplicando o Método LC. Relativamente ao comportamento do parâmetro \hat{k}_t , tal como no método anterior, este apresenta uma tendência decrescente, porém é o parâmetro em que existem mais diferenças associadas ao tipo de método aplicado. Tal como foi referido na secção 6.2, o ano 2020 não é considerado para a projeção, uma vez que é considerado um ano atípico, com valores muito mais elevados de mortes.

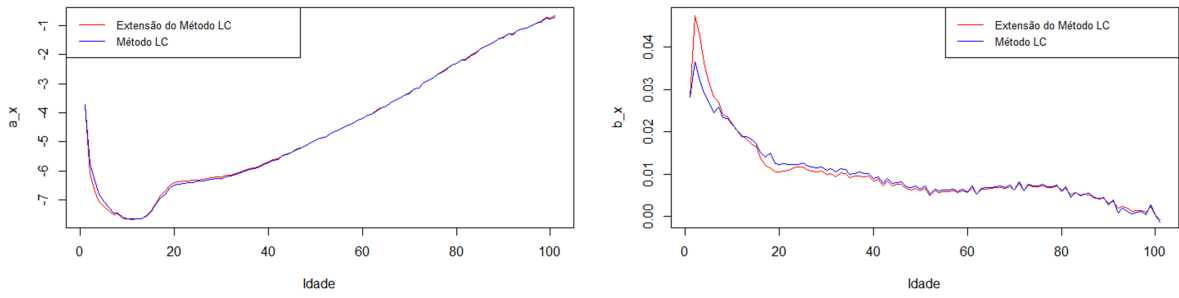


Figura 6.49: Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sexo Masculino.

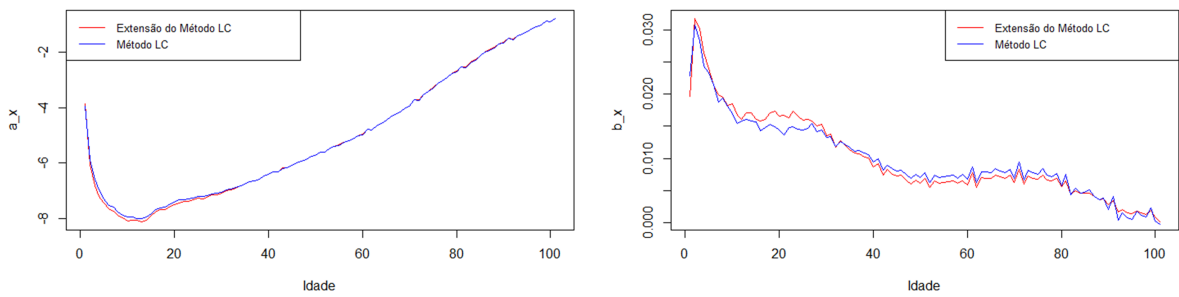


Figura 6.50: Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sexo Feminino.

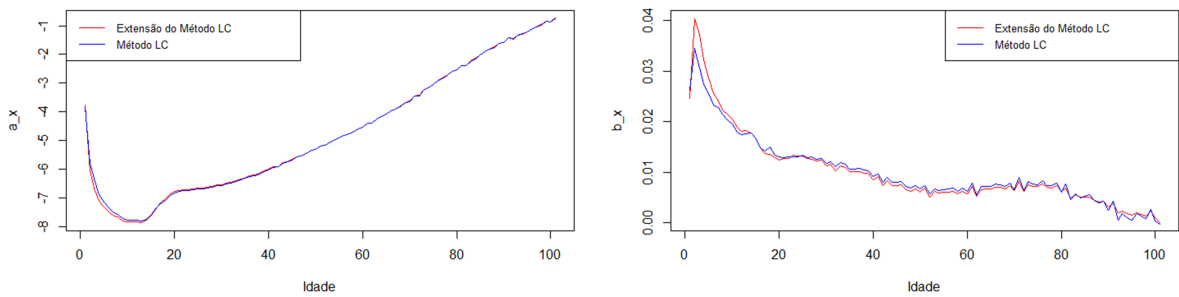


Figura 6.51: Estimativas dos parâmetros α_x e β_x nas idades 0-100 anos - Sem Distinção de Sexo.

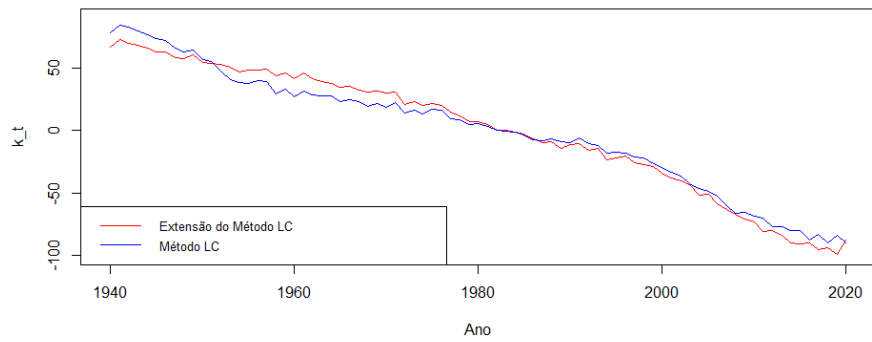


Figura 6.52: Estimativas do parâmetro k_t nos anos 1940-2020 - Sexo Masculino.

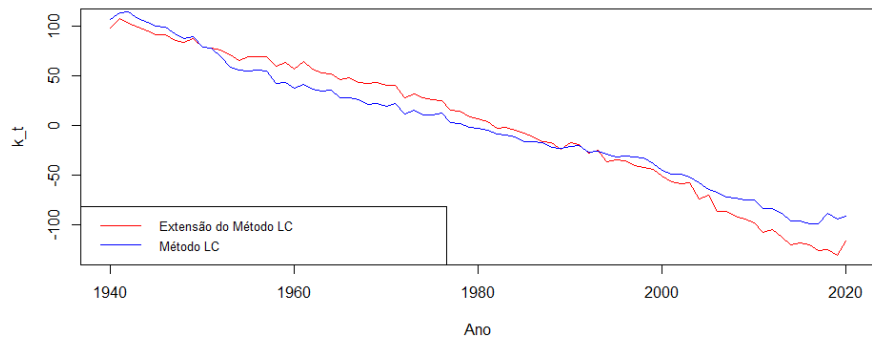


Figura 6.53: Estimativas do parâmetro k_t nos anos 1940-2020 - Sexo Feminino.

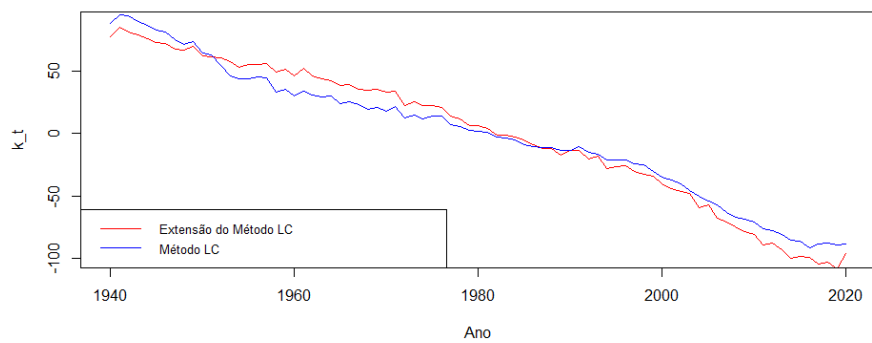


Figura 6.54: Estimativas do parâmetro k_t nos anos 1940-2020 - Sem Distinção de Sexo.

6.6 Modelação do Nível Geral de Mortalidade - Modelação ARIMA

Tal como no Método Lee-Carter, o parâmetro k_t necessita de ser modelado para posteriormente ser possível realizar projeções do mesmo. A série original também foi dividida em duas subséries, a série de treino que contém as observações dos anos 1940 a 2010 e a série de teste que é constituída pelas observações dos anos 2011 a 2019, de forma a realizar o processo de modelação e previsão desta série. Outra nota importante é que cada série foi analisada consoante o sexo, ou seja, para o sexo masculino, feminino e sem distinção de sexo.

6.6.1 Identificação do Modelo

A primeira subetapa para a identificação do modelo passa por perceber se as séries são estacionárias ou não. Nas Figuras 6.55, 6.56 e 6.57 estão representadas as séries da estimativa do parâmetro \hat{k}_t para cada sexo. Com base nas mesmas constata-se que as séries temporais parecem apresentar estacionariedade na variância e não na média, uma vez que todas apresentam uma tendência decrescente ao longo do tempo. Assim, as séries são não estacionárias.

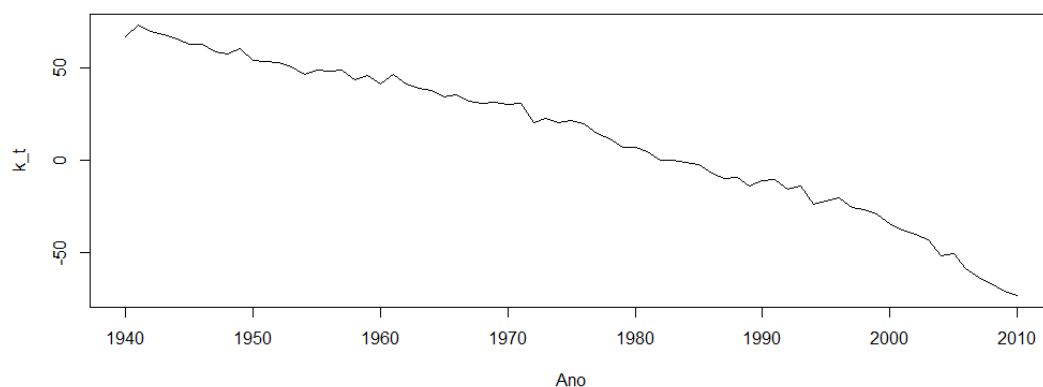


Figura 6.55: Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sexo Masculino.

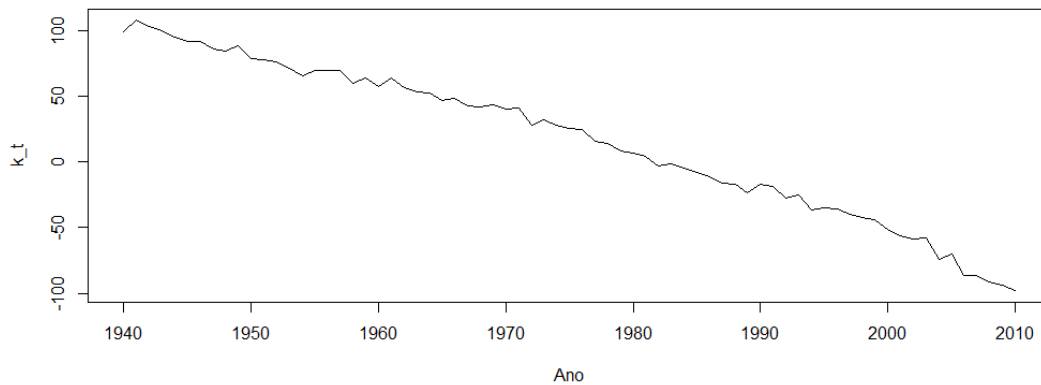


Figura 6.56: Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sexo Feminino.

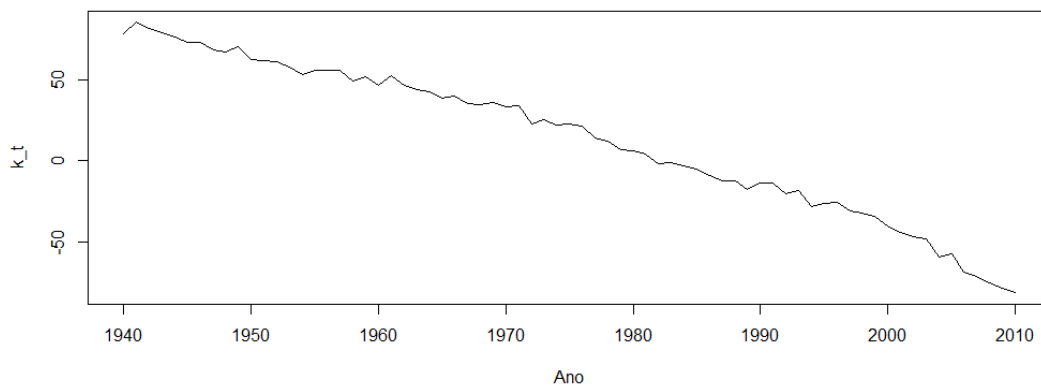


Figura 6.57: Série temporal do nível geral de mortalidade k_t no período de treino (1940-2010) - Sem Distinção de Sexo.

Para confirmar se as séries são realmente não estacionárias também é possível analisar a FAC empírica das mesmas e realizar os testes ADF e KPSS que testam a estacionariedade de uma série. Nas Figuras 6.58, 6.59 e 6.60 observa-se que a FAC empírica das três séries apresenta um decaimento lento para zero à medida que o *lag* aumenta, isto indica que as séries são não estacionárias. Para rejeitar a hipótese de não estacionariedade através dos testes ADF e KPSS é necessário que as estatísticas de teste sejam menores do que os valores críticos que, para um nível de significância de 5%, são -1,95 e 0,463, respetivamente. No teste ADF são obtidas as seguintes estatísticas de teste, 0,942 no caso do sexo masculino, 0,799 para o sexo feminino e 0,841 no caso sem distinção de sexo. Já para o teste KPSS as estatísticas de teste são 0,710 no sexo masculino, 0,712 no sexo feminino e 0,712 sem distinção de sexo. Uma vez que todas as estatísticas de teste são inferiores aos respetivos valores

críticos, comprova-se que as séries são não estacionárias, sendo que é necessário realizar a primeira diferenciação e, de seguida, testar novamente a estacionariedade das mesmas.

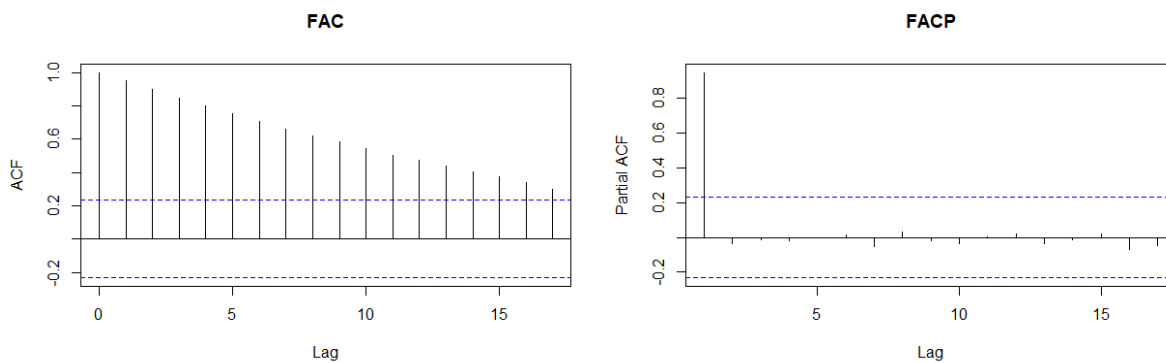


Figura 6.58: Função de autocorrelação e de autocorrelação parcial da série temporal do nível geral de mortalidade - Sexo Masculino.

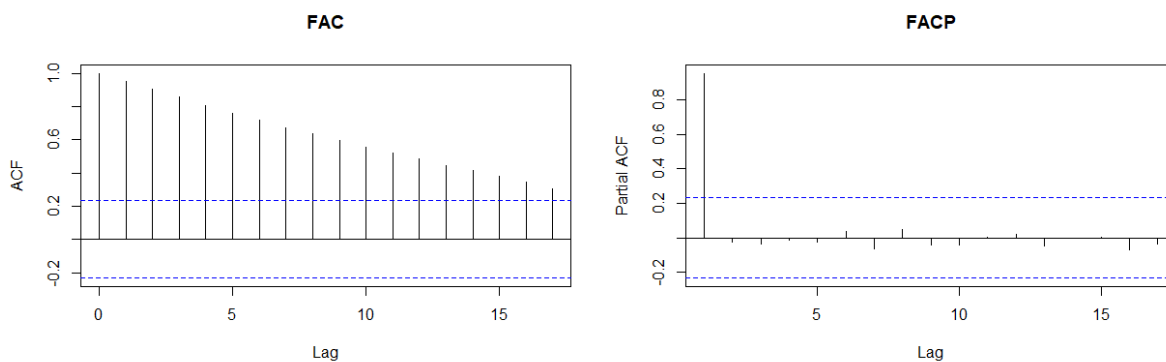


Figura 6.59: Função de autocorrelação e de autocorrelação parcial a série temporal do nível geral de mortalidade - Sexo Feminino.

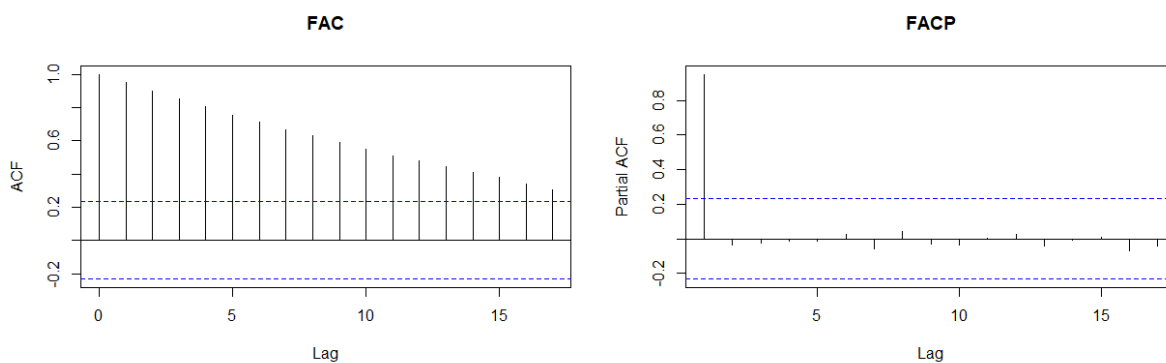


Figura 6.60: Função de autocorrelação e de autocorrelação parcial a série temporal do nível geral de mortalidade - Sem Distinção de Sexo.

Nas Figuras 6.61, 6.62 e 6.63 estão apresentadas as séries após a 1ª diferenciação e as respectivas FAC e FACP estimadas. Através da observação da representação gráfica dos resíduos, a média parece ainda não ser constante, sendo talvez necessário realizar outra diferenciação na série.

Para testar se realmente as séries já são estacionárias são utilizados novamente os testes ADF e KPSS. No primeiro teste as estatísticas de teste obtidas são 0,7337 no caso do sexo masculino, 0,4884 no sexo feminino e 0,5648 no caso sem distinção de sexo. No teste KPSS as estatísticas de teste são 0,063 (sexo masculino), 0,0412 (sexo feminino) e 0,0508 (sem distinção de sexo). Mediante as estatísticas de teste obtidas no teste ADF não se rejeita a hipótese nula de não estacionariedade em nenhum dos casos, porém a partir das estatísticas de teste obtidas no teste KPSS a hipótese nula de estacionariedade não é rejeitada em nenhum dos casos. Como foi referido na Subsecção 6.3.1, no caso dos dois testes apresentarem conclusões diferentes como neste caso, considera-se que a série é não estacionária, logo é necessário realizar mais uma diferenciação ($d=2$) em cada uma das séries.

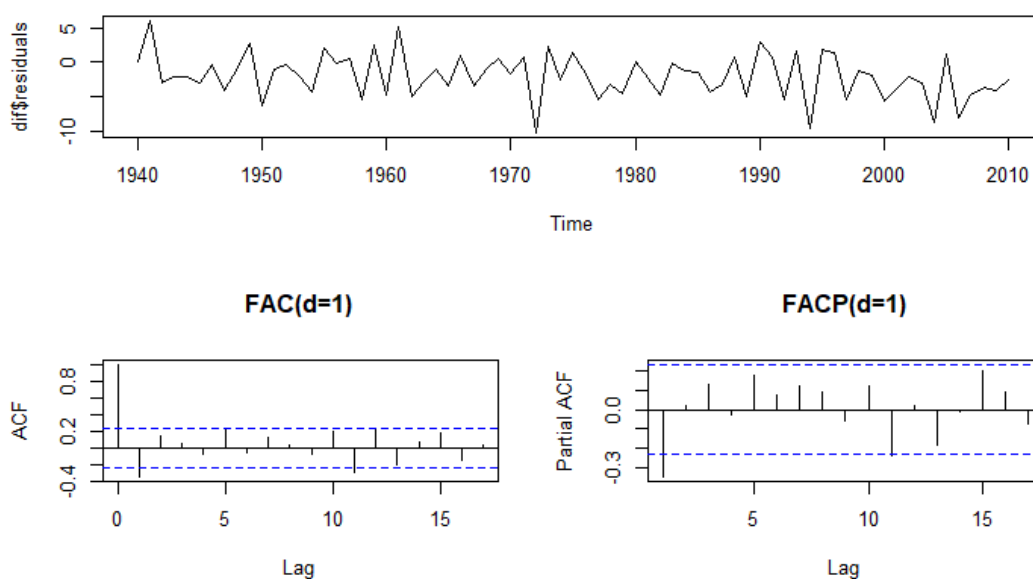


Figura 6.61: Série do nível geral de mortalidade k_t após a 1ª diferenciação ($d=1$) e respectivas FAC e FACP estimadas - Sexo Masculino.

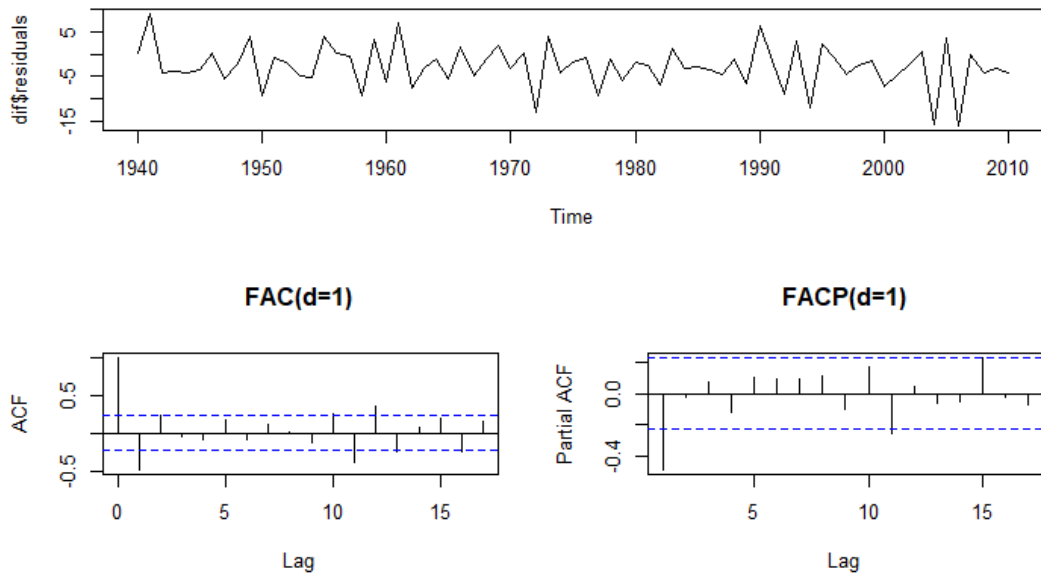


Figura 6.62: Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sexo Feminino.

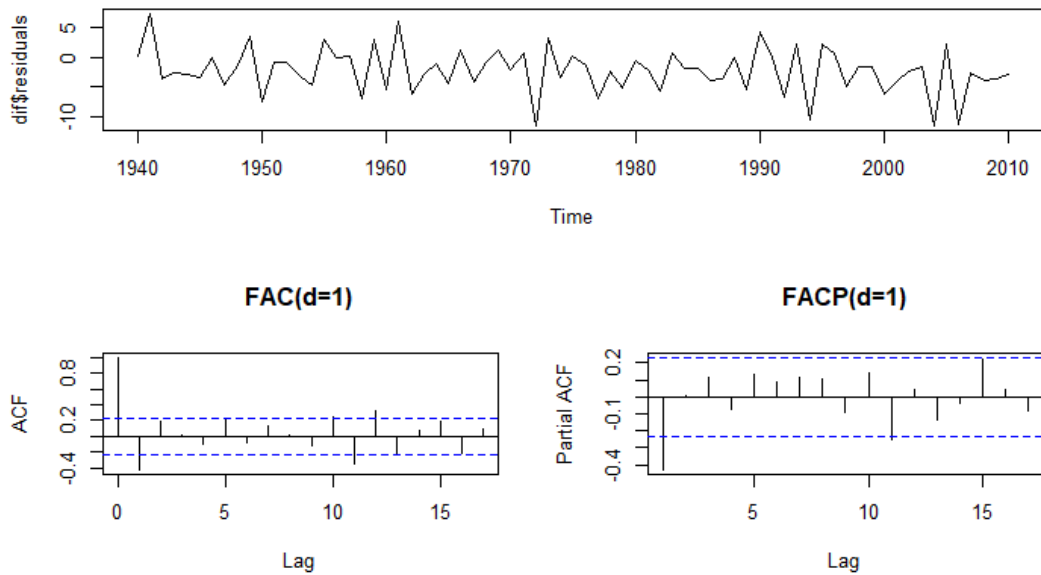


Figura 6.63: Série do nível geral de mortalidade k_t após a 1ª diferenciação (d=1) e respectivas FAC e FACP estimadas - Sem Distinção de Sexo.

Após realizada a 2ª diferenciação a cada série são obtidas as representações gráficas e a FAC e

FACP estimadas apresentadas nas Figuras 6.64, 6.65 e 6.66. Após a análise da representação gráfica dos resíduos de cada uma das séries, já é possível observar que as séries apresentam média constante em todos os casos, contudo é necessário realizar os testes ADF e KPSS para confirmar.

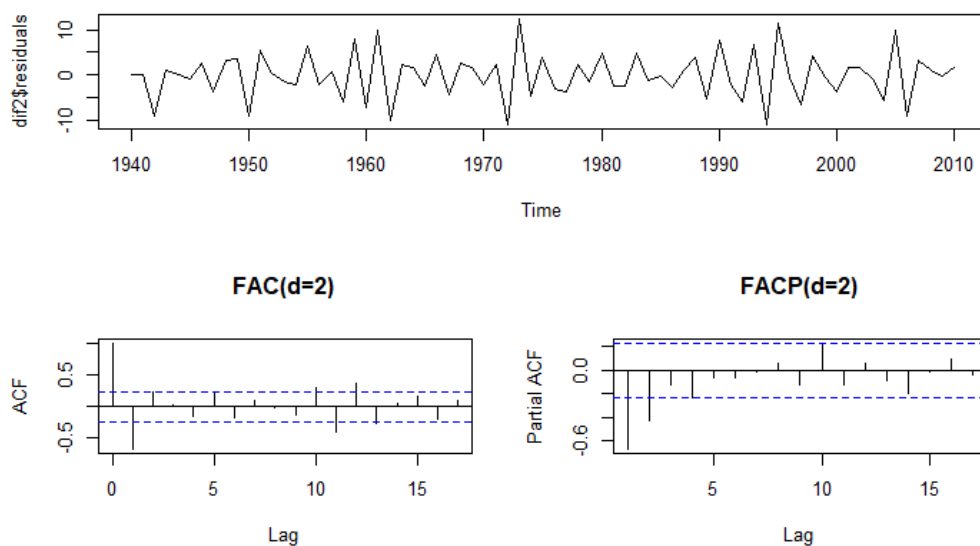


Figura 6.64: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respectivas FAC e FACP estimadas - Sexo Masculino.

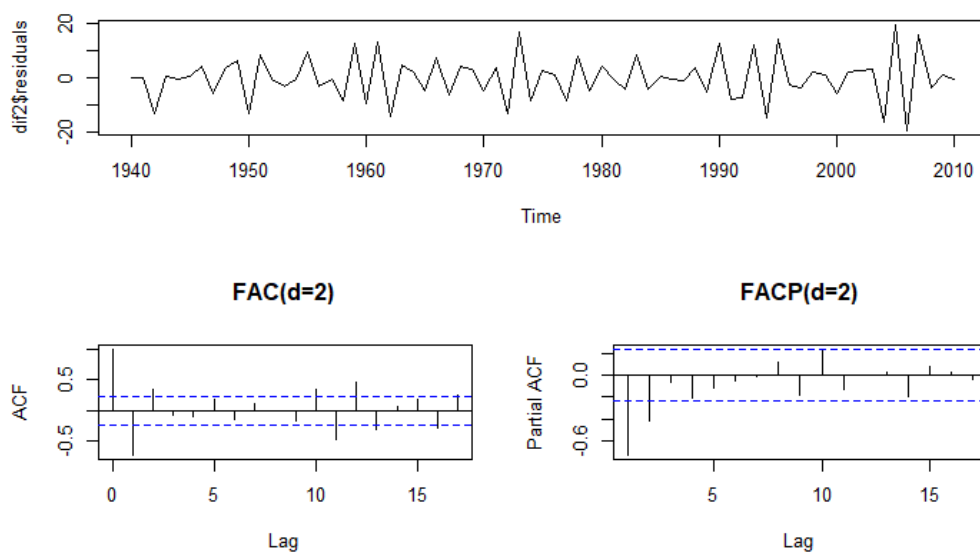


Figura 6.65: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respectivas FAC e FACP estimadas - Sexo Feminino.

Depois de aplicados os testes ADF e KPSS são obtidas as estatísticas de teste -2,795 e 0,192 no sexo masculino, -3,149 e 0,207 no sexo feminino e -2,962 e 0,206 no caso sem distinção de sexo, respetivamente. A partir destes resultados, com um nível de significância de 5%, conclui-se que em ambos os testes as séries são consideradas estacionárias. Como as séries já são estacionárias, é possível passar para a segunda subetapa, a escolha das ordens p e q do modelo ARIMA.

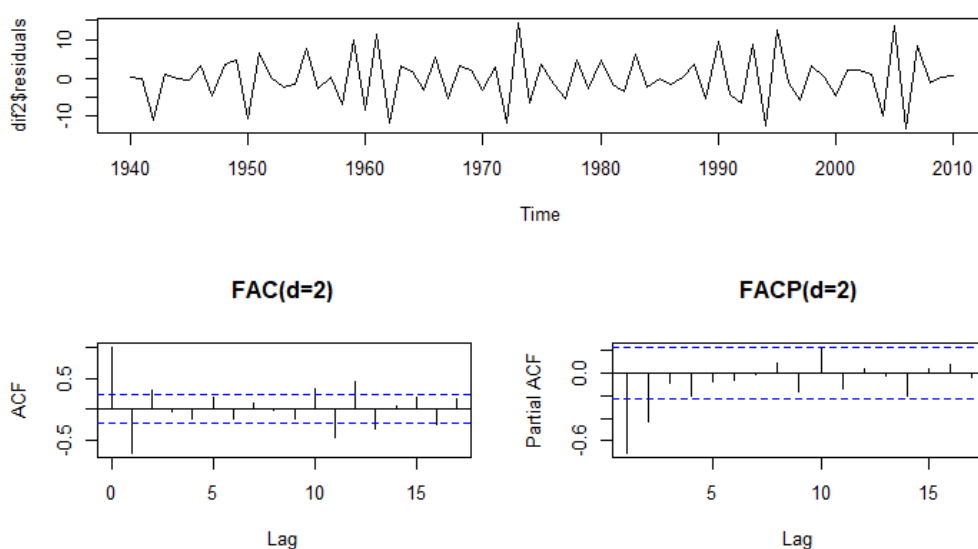


Figura 6.66: Série do nível geral de mortalidade k_t após a 2ª diferenciação ($d=2$) e respetivas FAC e FACP estimadas - Sem Distinção de Sexo.

Para a segunda subetapa foi utilizado, tal como na Subsecção 6.3.1, o algoritmo Hyndman-Khandakar (Hyndman & Athanasopoulos, 2021). Nas Tabelas 6.13, 6.14 e 6.15 estão apresentados os modelos considerados em cada fase do algoritmo e os respetivos valores de AIC e BIC. De notar que em todos os modelos a ordem d é igual a 2, uma vez que para todos os casos foi necessário realizar duas diferenciações para as séries se tornarem estacionárias. Na 1ª fase do algoritmo os modelos que apresentam menores valores de AIC e BIC são ARIMA(0,2,1) para os três casos. Após serem realizadas as variações deste modelo, ou seja, a 2ª fase do modelo, os modelos mais adequados são ARIMA(0,2,2) tanto no sexo masculino, como no feminino e no caso sem distinção de sexo.

Tabela 6.13: Identificação da ordem do modelo ARIMA - Sexo Masculino.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	426,33	428,56
	ARIMA(2,2,2)	0,07*	0,07*	-1,59	0,74	350,37	361,54
	ARIMA(0,2,1)	-	-	-0,95	-	364,13	368,60
	ARIMA(1,2,0)	-0,70	-	-	-	384,25	388,71
2ª Fase	ARIMA(0,2,2)	-	-	-1,39	0,49	348,81	355,52
	ARIMA(1,2,1)	-0,48	-	-0,90	-	350,78	357,48
	ARIMA(1,2,2)	0,05*	-	-1,57	0,72	348,57	357,50

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Tabela 6.14: Identificação da ordem do modelo ARIMA - Sexo Feminino.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	488,57	490,80
	ARIMA(2,2,2)	-1,20	-0,46	-0,38*	-0,49*	399,27	410,44
	ARIMA(0,2,1)	-	-	-0,98	-	420,21	424,68
	ARIMA(1,2,0)	-0,75	-	-	-	434,14	438,61
2ª Fase	ARIMA(0,2,2)	-	-	-1,57	0,69	397,16	403,86
	ARIMA(1,2,1)	-0,58	-	-0,93	-	397,20	403,91
	ARIMA(1,2,2)	-0,36*	-	-1,25	0,33*	397,49	406,43

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

Tabela 6.15: Identificação da ordem do modelo ARIMA - Sem Distinção de Sexo.

Fase do Algoritmo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
1ª Fase	ARIMA(0,2,0)	-	-	-	-	452,11	454,34
	ARIMA(2,2,2)	-1,18	-0,45	-0,33*	-0,50*	369,72	380,89
	ARIMA(0,2,1)	-	-	-0,96	-	386,42	390,89
	ARIMA(1,2,0)	-0,73	-	-	-	402,94	407,41
2ª Fase	ARIMA(0,2,2)	-	-	-1,55	0,69	365,97	372,67
	ARIMA(1,2,1)	-0,54	-	-0,91	-	367,83	374,53
	ARIMA(1,2,2)	-0,10*	-	-1,47	0,60*	367,78	376,71

* o coeficiente não é estatisticamente significativo, para um nível de significância $\alpha = 5\%$.

6.6.2 Avaliação do Modelo

Nesta etapa os modelos escolhidos anteriormente são testados para verificar se estes são realmente os mais adequados para cada caso. Para isso é necessário analisar a significância estatística dos parâmetros estimados e averiguar se os resíduos do modelo apresentam um comportamento semelhante ao de um ruído branco Gaussiano.

Na Tabela 6.16 estão apresentados os modelos ARIMA escolhidos na subsecção anterior e através desta verifica-se que os parâmetros de todos os modelos são estatisticamente significativos, com um nível de significância de 5%.

Tabela 6.16: Modelos ARIMA mais adequados para cada caso.

Sexo	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
Masculino	ARIMA(0,2,2)	-	-	-1,39	0,49	348,81	355,52
Feminino	ARIMA(0,2,2)	-	-	-1,57	0,69	397,16	403,86
Sem Distinção	ARIMA(0,2,2)	-	-	-1,55	0,69	365,97	372,67

Posto isto falta apenas analisar o comportamento dos resíduos de cada modelo, sendo que é necessário verificar se existe correlação entre os resíduos, através do teste de *Ljung-Box*, e se estes apresentam valor médio nulo, por aplicação do teste *t*.

Antes de mais, nas Figuras 6.67, 6.68 e 6.69 estão apresentadas as séries dos resíduos, o respetivo

histograma e a FAC e FACP de cada sexo. Em relação à distribuição dos resíduos, através dos histogramas constata-se que nos três casos estes apresentam uma distribuição assimétrica à esquerda, tal como aconteceu nas séries expostas na Subsecção 6.3.2, contudo esta assimetria não é significativa, sendo que, através do teste *Lilliefors*, a hipótese de normalidade não foi rejeitada em nenhum dos casos, tendo como valores de prova 0,962 no caso do sexo masculino, 0,974 no sexo feminino e 0,936 no caso sem distinção de sexo. Visto que a normalidade não foi rejeitada, é possível testar, através do teste *t*, se os resíduos apresentam valor médio nulo. Depois de aplicado este teste são obtidos os valores de prova 0,070, 0,175 e 0,176 nos casos sexo masculino, feminino e sem distinção de sexo, respetivamente, ou seja, não se rejeita a hipótese de média nula em nenhum dos casos, com um nível de significância de 5%. Assim sendo, falta apenas averiguar se os resíduos são ou não correlacionados. Para isso é utilizado o teste *Ljung-Box* que tem como valores de prova 0,787 (sexo masculino), 0,475 (sexo feminino) e 0,648 (sem distinção de sexo), logo em nenhum dos casos os resíduos são correlacionados. Depois de realizados todos os testes necessários conclui-se que os resíduos dos modelos apresentados anteriormente para cada caso exibem um comportamento análogo ao de um ruído branco Gaussiano, ou seja, são considerados modelos adequados.

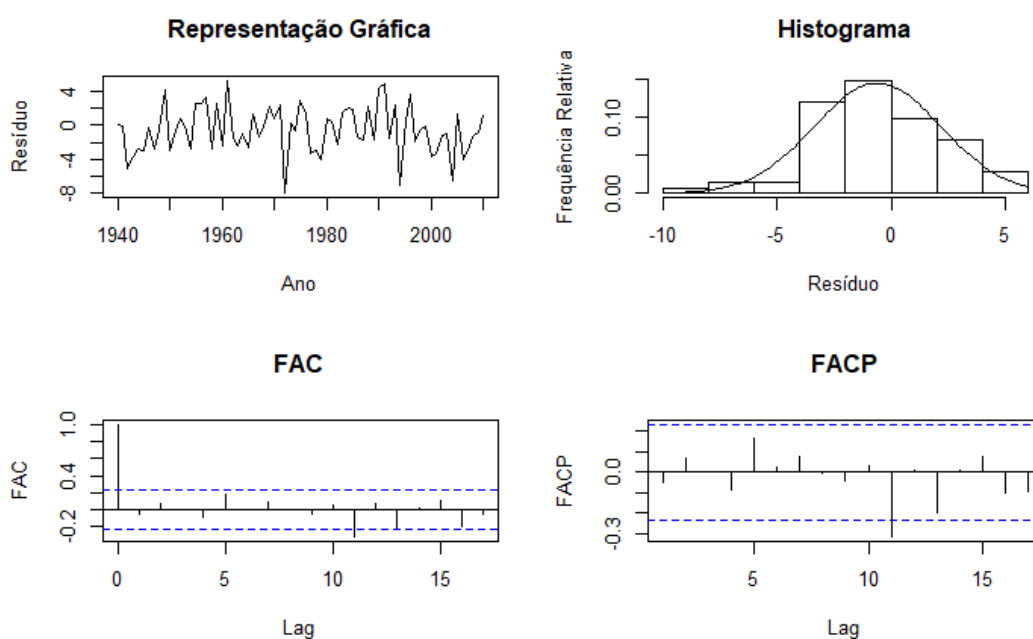


Figura 6.67: Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.

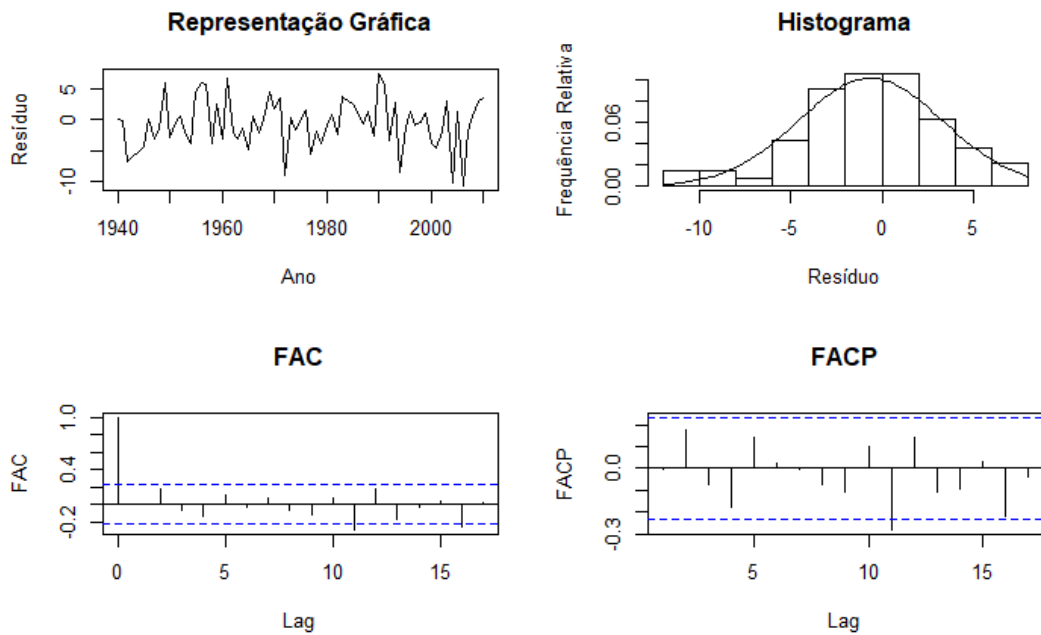


Figura 6.68: Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.

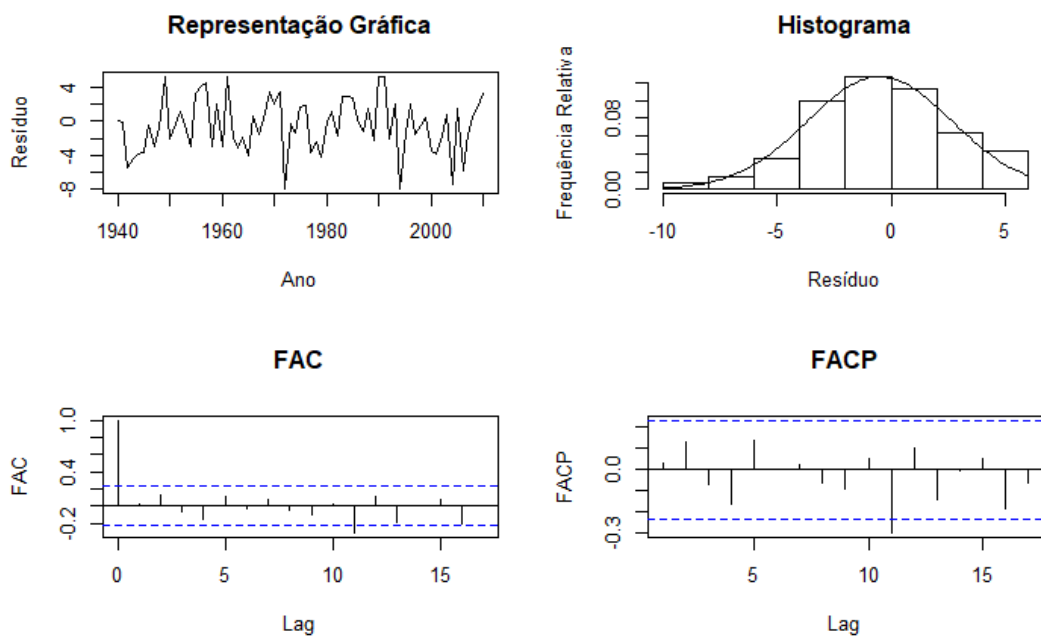


Figura 6.69: Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.

Depois de realizada a avaliação do modelo e concluindo-se que os modelos ARIMA escolhidos em

cada um dos casos são adequados, torna-se plausível a realização de estimativas e previsões para um período de tempo. Neste caso são efetuadas estimativas pontuais para o período de treino (1940-2010) e previsões pontuais e intervalares para o período de teste (2011-2019), sendo que estas podem ser observadas nas Figuras 6.70, 6.71 e 6.72, para cada um dos casos. Através das figuras compreende-se que a qualidade preditiva dos modelos é superior na série de treino do que na série de teste. Isto pode ser comprovado através do cálculo de medidas de avaliação nos dois períodos distintos, sendo que estas estão apresentadas na Tabela 6.17. Através dos valores obtidos comprova-se o que foi dito anteriormente, uma vez que os valores das medidas de avaliação são mais elevados nas séries de teste do que nas de treino. Ainda que estas medidas sejam elevadas na série de teste, todas as 9 observações de cada série encontram-se no intervalo de previsão, isto é, a taxa de cobertura é igual a 100% nos três casos.

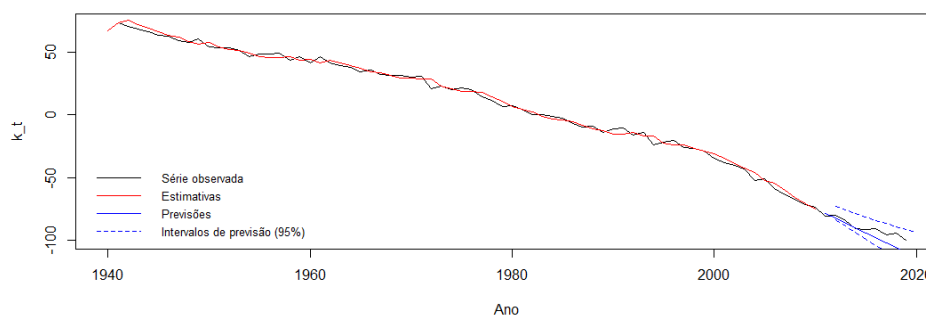


Figura 6.70: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.

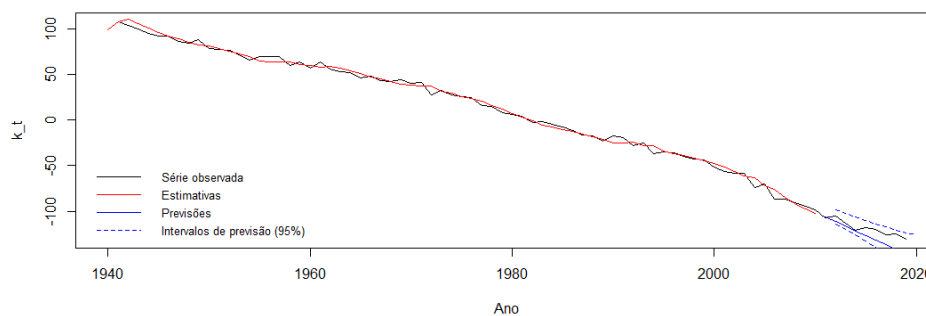


Figura 6.71: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.

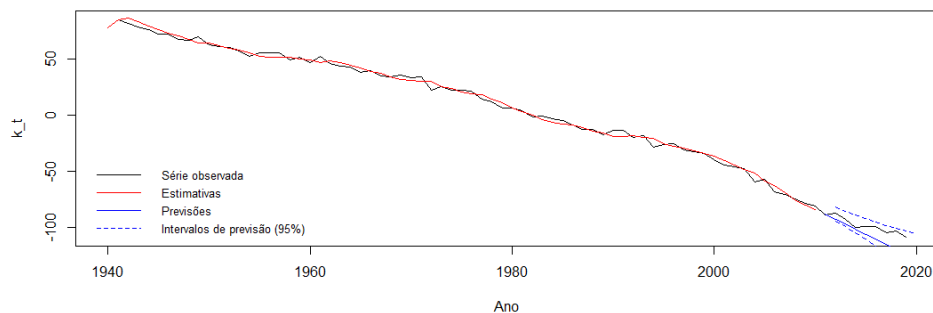


Figura 6.72: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.

Tabela 6.17: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA.

Sexo	EQM	REQM	EPAM	EAM	U-Theil	
Masculino	7,872	2,806	0,702	2,256	2,018	
Feminino	15,718	3,965	0,132	3,105	0,783	Série de Treino
Sem Distinção	10,027	3,167	0,160	2,556	1,256	
Masculino	39,465	6,282	0,055	5,065	1,578	
Feminino	103,118	10,155	0,069	8,411	2,037	Série de Teste
Sem Distinção	92,248	9,605	0,079	8,042	2,291	

6.7 Modelação do Nível Geral de Mortalidade - Método de Holt

Uma vez que também não é possível reconhecer sazonalidade nas séries do nível geral de mortalidade calculadas através da Extensão do Método LC, o método de alisamento exponencial mais adequado é novamente o Método Linear de Holt (ou Método de Holt). Na Tabela 6.18 estão expostas as estimativas iniciais para o nível e declive e as estimativas das constantes de alisamento. Comparativamente com os resultados obtidos na Tabela 6.11, os valores da constante de alisamento associada ao nível neste caso, $\hat{\alpha}$, são bem menores do que os observados na Secção 6.4 e, como estes se encontram mais perto de 0 do que de 1, constata-se que é dado mais peso às observações mais distantes. Em relação aos valores da constante de alisamento associada ao declive, $\hat{\beta}$, estes continuam a ser valores baixos e próximos de 0, logo a tendência quase não muda ao longo do tempo.

Tabela 6.18: Estimativas iniciais para o nível e declive e estimativas das constantes de alisamento, resultantes da aplicação do Método de Holt ao parâmetro k_t produzido a partir da Extensão do Método Lee-Carter.

Sexo	Estimativas			
	$\hat{\alpha}$	$\hat{\beta}$	\hat{l}_1	\hat{b}_1
Homens	0,272	0,158	73,839	-1,125
Mulheres	0,280	0,116	108,229	-2,199
Sem Distinção	0,284	0,140	85,736	-1,540

Nas Figuras 6.73, 6.74 e 6.75 estão representadas as estimativas pontuais e as previsões pontuais e intervalares para cada um dos casos. Após a observação das mesmas, repara-se que a qualidade preditiva é superior no período de treino do que no período de teste. Esta premissa pode ser comprovada a partir do cálculo de medidas de avaliação para cada período, sendo que estas estão presentes na Tabela 6.19. É notório que existe uma grande diferença nos valores das medidas calculadas nos dois períodos, na medida em que estes valores são bastante superiores no período de teste, comparativamente com os do período de treino. Não obstante, apesar destes resultados, as observações das três séries de teste estão inseridos nos respetivos intervalos de previsão, apresentando uma taxa de cobertura de 100%.

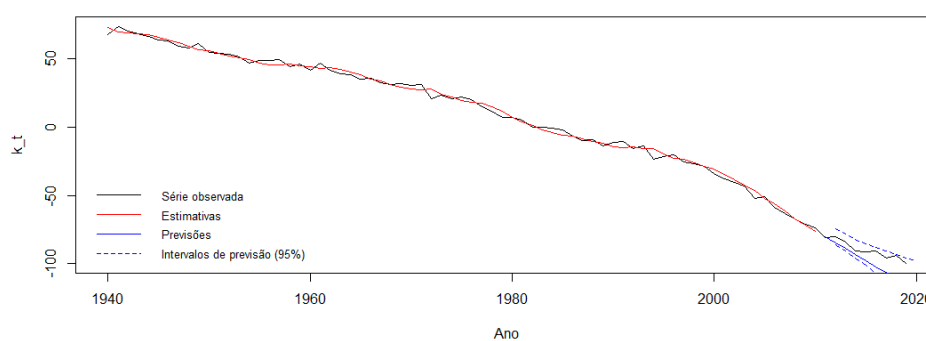


Figura 6.73: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.

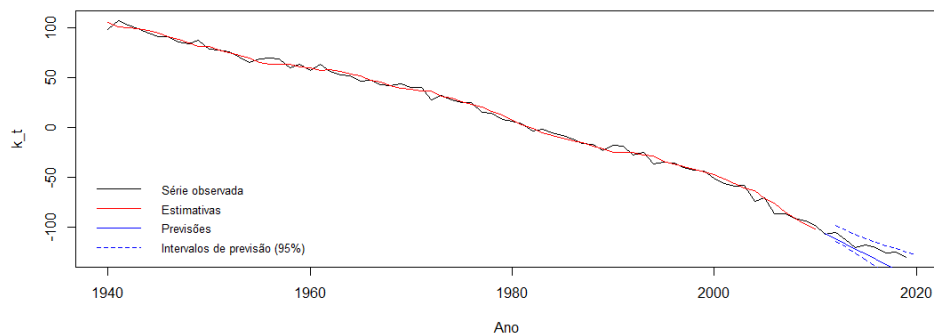


Figura 6.74: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.

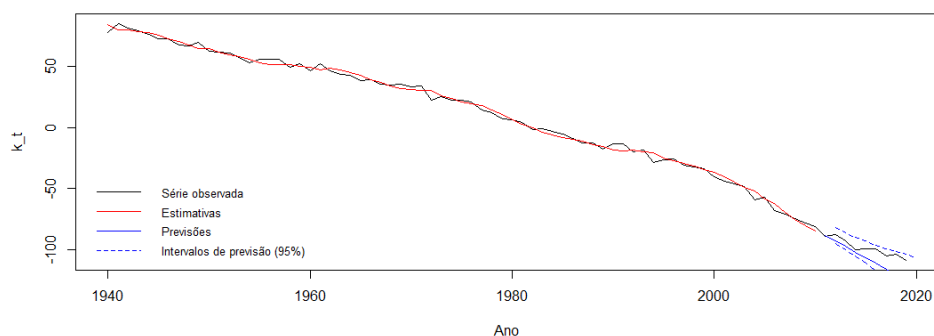


Figura 6.75: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do Método de Holt, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.

Tabela 6.19: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt.

Sexo	EQM	REQM	EPAM	EAM	U-Theil	
Masculino	7,573	2,752	0,725	2,239	3,334	
Feminino	15,696	3,962	0,134	3,126	0,817	Período de Treino
Sem Distinção	10,014	3,164	0,162	2,557	1,309	
Masculino	96,111	9,804	0,089	8,224	2,510	
Feminino	119,475	10,930	0,075	9,116	2,197	Período de Teste
Sem Distinção	100,895	10,045	0,083	8,441	2,399	

6.8 Método Lee-Carter vs. Extensão do Método Lee-Carter

Nesta secção são comparados os métodos de modelação, ARIMA e Método de Holt, utilizados para modelar o nível geral de mortalidade calculado a partir do Método Lee-Carter e da Extensão do Método Lee-Carter (Modelo Poisson), com a finalidade de perceber qual é o mais adequado para a situação em estudo.

6.8.1 Modelação ARIMA

Para ser feita a comparação entre os modelos escolhidos nas secções anteriores, é necessário utilizar alguns critérios como os valores de AIC e BIC dos modelos e as medidas de avaliação calculadas tanto na série de treino como na série de teste, que são o EQM, REQM, EPAM, EAM e a estatística de U-Theil.

No caso do sexo masculino foi possível encontrar um modelo através do algoritmo Hyndman-Khandakar (Hyndman & Athanasopoulos, 2021) que cumprisse todos os pressupostos pretendidos em ambos os métodos. Na Tabela 6.20 estão apresentados os modelos mais adequados para cada método e os respetivos parâmetros e valores de AIC e BIC e na Tabela 6.21 estão as cinco medidas de avaliação calculadas para a série de treino e para a série de teste em ambos os métodos.

Tabela 6.20: Modelos ARIMA mais adequados para cada método - Sexo Masculino.

Método	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$	AIC	BIC
Lee-Carter	ARIMA(1,2,1)	-0,34	-	-0,88	-	367,65	374,36
Extensão de Lee-Carter	ARIMA(0,2,2)	-	-	-1,39	0,49	348,81	355,52

Tabela 6.21: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Masculino.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	10,427	3,229	0,315	2,617	0,600	Série de Treino
Extensão de Lee-Carter	7,872	2,806	0,702	2,256	2,018	
Lee-Carter	39,199	6,261	0,055	4,579	1,111	Série de Teste
Extensão de Lee-Carter	39,465	6,282	0,055	5,065	1,578	

Através dos valores de AIC e BIC presentes na primeira tabela, constata-se que o modelo identificado

a partir da série temporal obtida pela Extensão do Método LC apresenta valores menores nestas duas medidas, sendo por isso considerado como mais adequado para modelação, comparativamente com o modelo considerado para a série temporal obtida através do Método LC. Relativamente às medidas de avaliação calculadas nas séries de treino e de teste de ambos os métodos, constata-se o que foi referido anteriormente, uma vez que se verifica que o modelo que explica melhor o comportamento dos dados (série de treino) é o resultante da Extensão do Método LC. Porém, em relação à previsão (série de teste), este é o modelo que apresenta piores resultados. Este resultado é bastante comum dado que o modelo que melhor se ajusta aos dados históricos pode não ser o melhor para gerar previsões (Chatfield, 2000).

Como o objetivo deste trabalho é encontrar o modelo mais adequado para realizar previsões para anos futuros, optou-se por testar o comportamento das previsões de todos os modelos que apresentavam coeficientes estatisticamente significativos, utilizando tanto a série de nível geral de mortalidade obtida através do Método Lee-Carter como a obtida através da Extensão do Método Lee-Carter. Para isso, as previsões obtidas através de cada modelo são comparadas com as observações originais da série de teste, onde são calculadas as respetivas medidas de avaliação, de forma a encontrar o modelo que apresenta menores valores das mesmas. Na Tabela 6.22 são apresentadas as medidas de avaliação obtidas em cada modelo.

Tabela 6.22: Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Masculino.

Modelo	EQM	REQM	EPAM	EAM	U-Theil	Método
ARIMA(0,2,0)	16,085	4,011	0,039	3,520	1,000	
ARIMA(1,2,0)	7,638	2,764	0,024	2,155	0,646	
ARIMA(0,2,1)	9,227	3,038	0,028	2,515	0,725	Extensão do Método LC
ARIMA(0,2,2)	39,465	6,282	0,055	5,065	1,578	
ARIMA(1,2,1)	12,874	3,588	0,031	2,803	0,864	
ARIMA(0,2,0)	55,415	7,444	0,068	5,674	1,326	
ARIMA(1,2,0)	26,018	5,101	0,051	4,257	1,006	
ARIMA(0,2,1)	12,086	3,476	0,035	2,889	0,684	Método LC
ARIMA(0,2,2)	76,283	8,734	0,083	6,889	1,561	
ARIMA(1,2,1)	39,199	6,261	0,055	4,579	1,111	

Mediante os valores presentes na tabela anterior constata-se que os modelos que apresentam menores valores são resultantes da série de nível geral de mortalidade obtida através da Extensão do Método LC (Modelo de Poisson), sendo que o modelo que apresenta melhores resultados é o ARIMA(1,2,0), porém é ainda necessário avaliar se os resíduos produzidos pelo modelo verificam os pressupostos pretendidos. Para isso, tal como nas Subsecções 6.3.2 e 6.6.2, é necessário testar a independência dos resíduos e se estes apresentam valor médio nulo e variância constante. Na Figura 6.76 observa-se que os resíduos apresentam uma distribuição ligeiramente enviesada à esquerda, sendo que o teste *Lilliefors* será importante para testar se os resíduos apresentam uma distribuição Normal ou não. Este teste obteve como valor prova 0,277 logo, com um nível de significância de 5%, a hipótese de normalidade não é rejeitada. Em relação ao valor médio nulo, o valor de prova obtido no teste t é 0,681 e, quanto à independência dos resíduos, utilizando o teste *Ljung-Box*, o valor de prova é 0,030 logo, apesar da hipótese de valor médio nulo não ser rejeitada, verifica-se que há correlação entre os resíduos, pelo que este modelo não é considerado adequado.

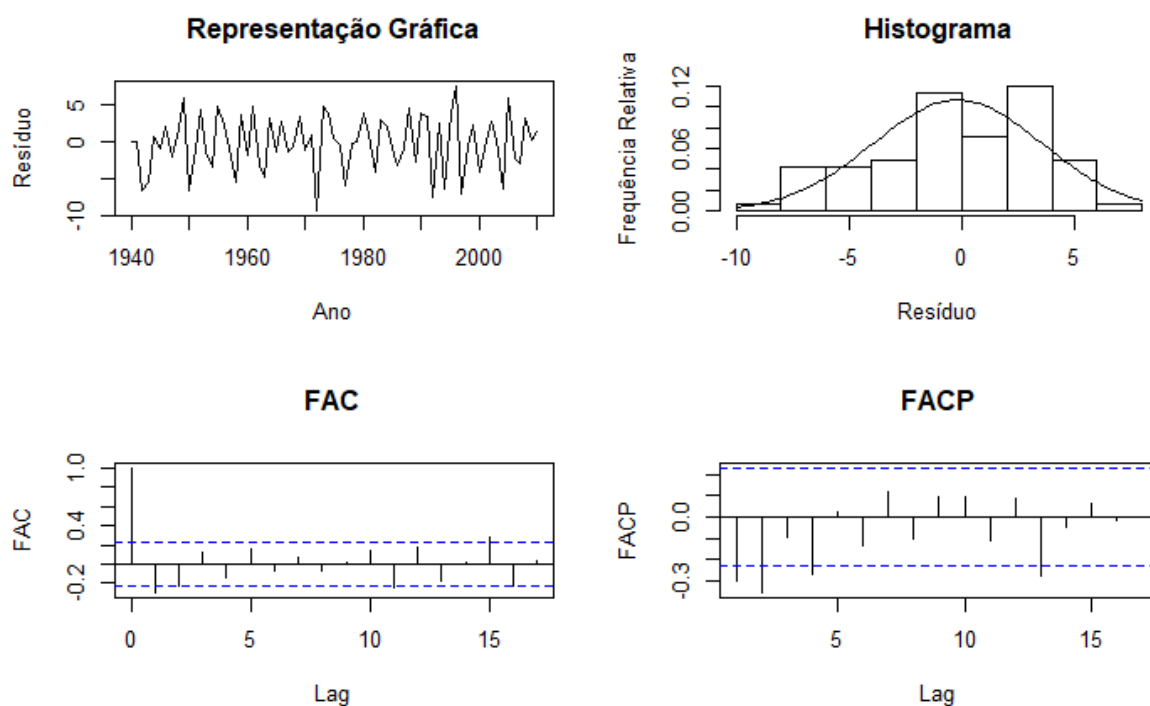


Figura 6.76: Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.

Posto isto é considerado o modelo alternativo, ou seja, o modelo que apresenta os segundos menores valores de medidas de avaliação que neste caso é o modelo ARIMA(0,2,1). Tal como no modelo anterior,

é indispensável analisar o comportamento dos resíduos produzidos pelo mesmo. Mediante o histograma da Figura 6.77 repara-se que, tal como no modelo anterior, os resíduos apresentam uma distribuição relativamente enviesada à esquerda, porém o valor de prova obtido no teste *Lilliefors* é 0,884, logo a hipótese de normalidade não é rejeitada. O teste *t* obteve como valor de prova 0,078 e o teste *Ljung-Box* obteve 0,005, por isso, mais uma vez, com um nível de significância de 5%, a hipótese de valor médio não é rejeitada mas constata-se a existência de correlação entre os resíduos. Deste modo é necessário recorrer ao modelo seguinte que apresenta menores medidas de avaliação, ou seja, o modelo ARIMA(1,2,1).

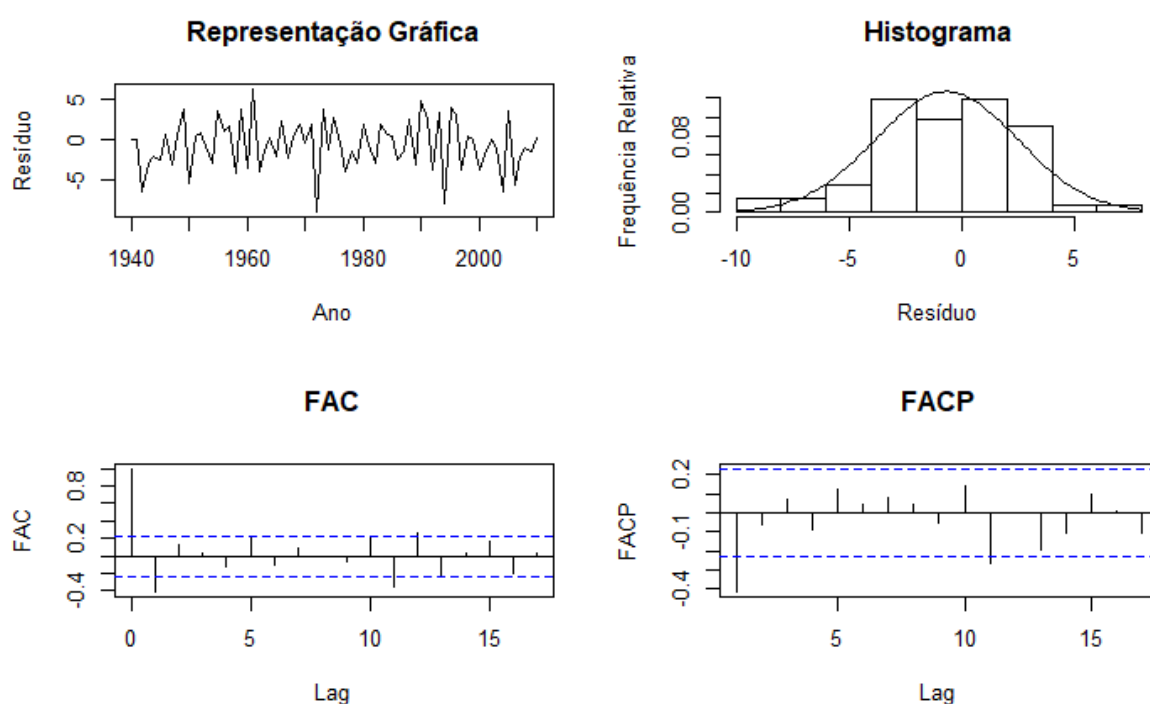


Figura 6.77: Série dos resíduos da série do nível geral de mortalidade k_t após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Masculino.

Os gráficos necessários para a análise dos resíduos do modelo ARIMA(1,2,1) estão representados na Figura 6.78 e a partir do histograma presente nesta verifica-se mais uma vez que os resíduos apresentam uma distribuição assimétrica à esquerda, porém esta não é suficiente para que a hipótese de normalidade seja rejeitada no teste *Lilliefors* (valor de prova igual a 0,971). A hipótese de valor médio nulo também não é rejeitada, sendo que o valor de prova obtido no teste *t* é 0,064 e através da análise gráfica constata-se que os resíduos oscilam uniformemente em torno de zero, o que leva a crer que a variância é constante. Por último, a hipótese de independência entre os resíduos, testada pelo teste *Ljung-Box*, também não é rejeitada, apresentando como valor de prova 0,823. Sendo assim, com um nível de significância de 5%,

comprova-se que os resíduos do modelo ARIMA(1,2,1) apresentam um comportamento semelhante ao de um ruído branco Gaussiano e, por isso, este modelo é adequado para realizar previsões. Na Figura 6.79 encontra-se a representação gráfica das estimativas pontuais no período de treino e as previsões pontuais e intervalares do período de teste utilizando o modelo testado anteriormente e, através desta, observa-se que as 9 observações originais que constituem a série de teste encontram-se compreendidas no intervalo de previsão, isto é, a taxa de cobertura é de 100%.

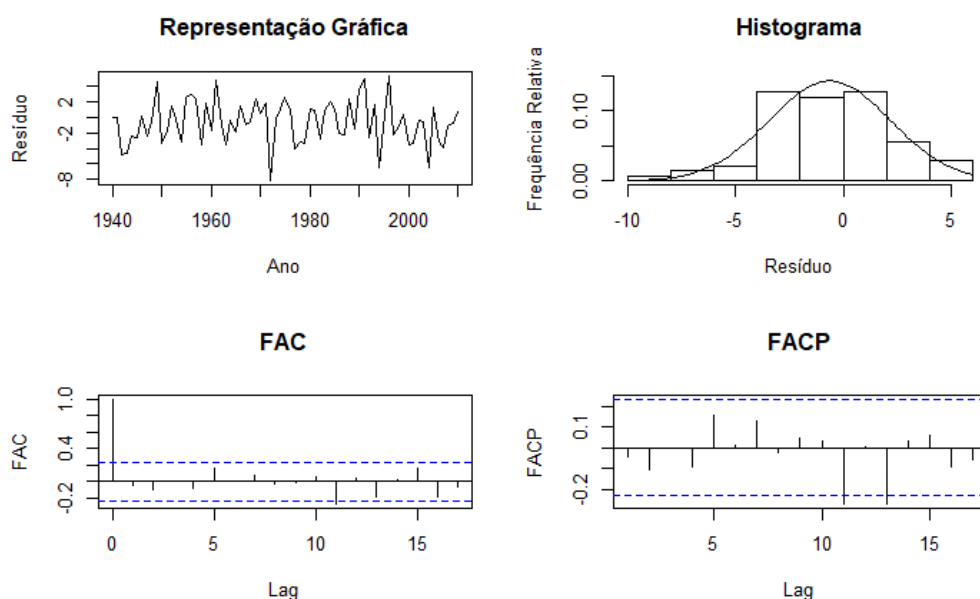


Figura 6.78: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP - Sexo Masculino.

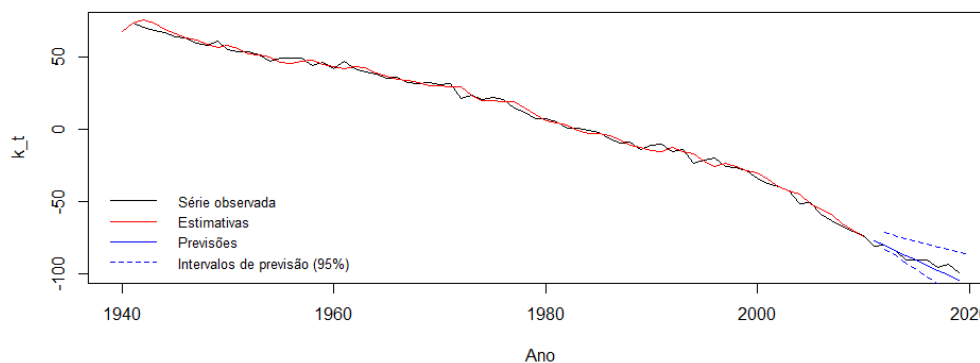


Figura 6.79: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Masculino.

Relativamente ao sexo feminino, na Tabela 6.23 estão descritos os modelos ARIMA considerados adequados para modelação, utilizando cada um dos métodos, e os respetivos valores de AIC e BIC. Atendendo aos valores de AIC e BIC presentes na tabela constata-se que o modelo identificado para a série temporal obtida através do Método Lee-Carter apresenta melhores valores de AIC e BIC, isto é, valores mais reduzidos, por essa razão é considerado como mais adequado para modelação do que o modelo identificado para a série temporal obtida através da Extensão do Método LC. Na Tabela 6.24 são apresentadas as medidas de avaliação calculadas na série de treino e na série de teste de ambos os modelos, com o intuito de verificar se o modelo escolhido explica tão bem o comportamento dos dados na série de treino como realiza as previsões na série de teste. As medidas declaradas na tabela indicam que o modelo da série temporal obtida através da aplicação do Método Lee-Carter é o mais preciso tanto para a série de treino como para a série de teste.

Tabela 6.23: Modelos ARIMA mais adequados para cada método - Sexo Feminino.

Método	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	AIC	BIC
Lee-Carter	ARIMA(2,2,3)	-1,63	-0,98	-	-	0,54	-0,60	-0,94	377,60	391,01
Extensão de LC	ARIMA(0,2,2)	-	-	-	-	-1,57	0,69	-	397,16	403,86

Tabela 6.24: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Feminino.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	10,310	3,211	0,116	2,529	0,742	Série de Treino
Extensão de Lee-Carter	15,718	3,965	0,132	3,105	0,783	
Lee-Carter	47,209	6,871	0,070	6,500	1,321	Série de Teste
Extensão de Lee-Carter	103,118	10,155	0,069	8,411	2,037	

Porém, tal como foi referido anteriormente, o modelo mais adequado para ajustar os dados pode não ser o mais adequado para prever observações futuras. Por essa razão, também no sexo feminino, é testado se existe um modelo mais adequado para realizar previsões, utilizando tanto a série de nível geral de mortalidade obtida através do Método Lee-Carter como a obtida através da Extensão do Método Lee-Carter. Em ambos os casos foram testados apenas os modelos que apresentavam coeficientes estatisticamente significativos e estes foram comparados através do cálculo de medidas de avaliação

entre as previsões obtidas e as observações originais (Tabela 6.25).

Tabela 6.25: Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sexo Feminino.

Modelo	EQM	REQM	EPAM	EAM	U-Theil	Método
ARIMA(0,2,0)	14,532	3,812	0,027	3,222	0,705	
ARIMA(1,2,0)	13,106	3,620	0,023	2,690	0,667	
ARIMA(0,2,1)	31,792	5,638	0,043	5,083	1,118	Extensão do Método LC
ARIMA(0,2,2)	103,118	10,155	0,069	8,411	2,037	
ARIMA(1,2,1)	12,370	3,517	0,023	2,713	0,645	
ARIMA(2,2,3)	47,209	6,871	0,070	6,500	1,321	
ARIMA(4,2,1)	36,445	6,037	0,054	5,008	1,172	Método LC

Através da análise da Tabela 6.25 observa-se que os modelos aplicados à série do nível geral de mortalidade calculado a partir da Extensão do Método LC apresentam, no geral, uma melhor qualidade preditiva comparativamente com os modelos da série temporal obtida através do Método Lee-Carter, sendo que o modelo que apresenta menores medidas de avaliação é o modelo ARIMA(1,2,1). Posto isto falta apenas averiguar se os resíduos do modelo verificam os pressupostos pretendidos. Mediante a Figura 6.80 verifica-se que os resíduos do modelo apresentam uma distribuição ligeiramente assimétrica à esquerda porém, através do teste de *Lilliefors* a hipótese de normalidade não é rejeitada com um valor de prova igual a 0,136. Relativamente à independência dos resíduos e ao valor médio, o primeiro pressuposto é testado através do teste de *Ljung-Box*, onde se conclui, com um valor de prova igual a 0,721, que os resíduos são não correlacionados, já o segundo pressuposto é testado a partir do teste *t*, obtendo-se um valor de prova igual a 0,072, ou seja, os resíduos apresentam valor médio nulo. O último pressuposto que falta averiguar é a variância constante e, através do gráfico dos resíduos observa-se que estes oscilam uniformemente em torno de zero, o que leva a crer que a variância é constante. Assim sendo, pode-se considerar que este modelo é adequado aos dados. Posto isto, na Figura 6.81 observa-se a representação gráfica das das estimativas pontuais no período de treino e as previsões pontuais e intervalares do período de teste utilizando o modelo mencionado anteriormente e constata-se que a taxa de cobertura é igual a 100%, o que indica que a totalidade das observações originais da série de teste estão incluídas no intervalo de previsão.

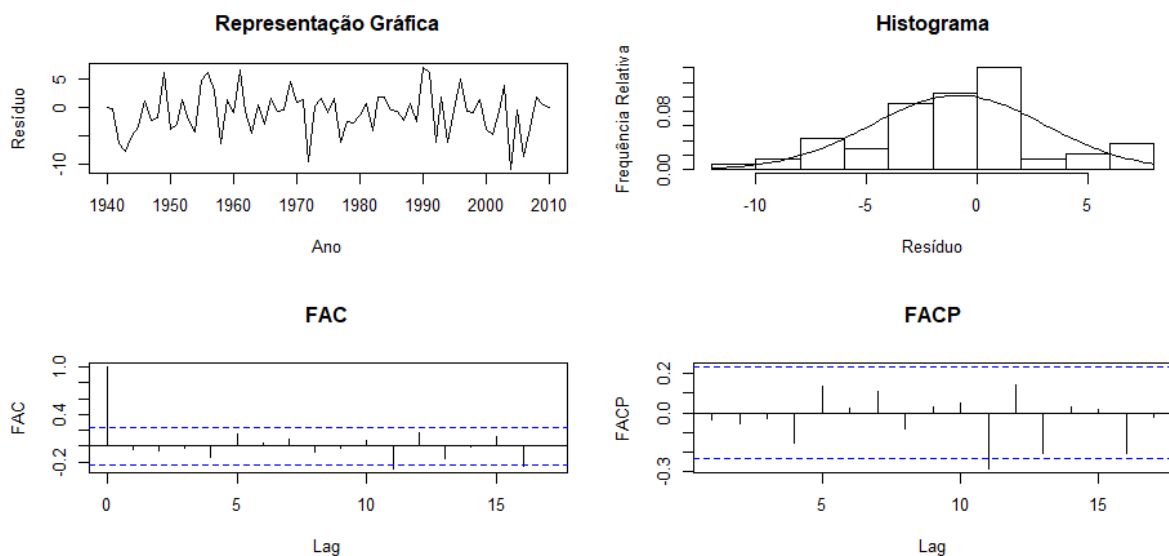


Figura 6.80: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sexo Feminino.

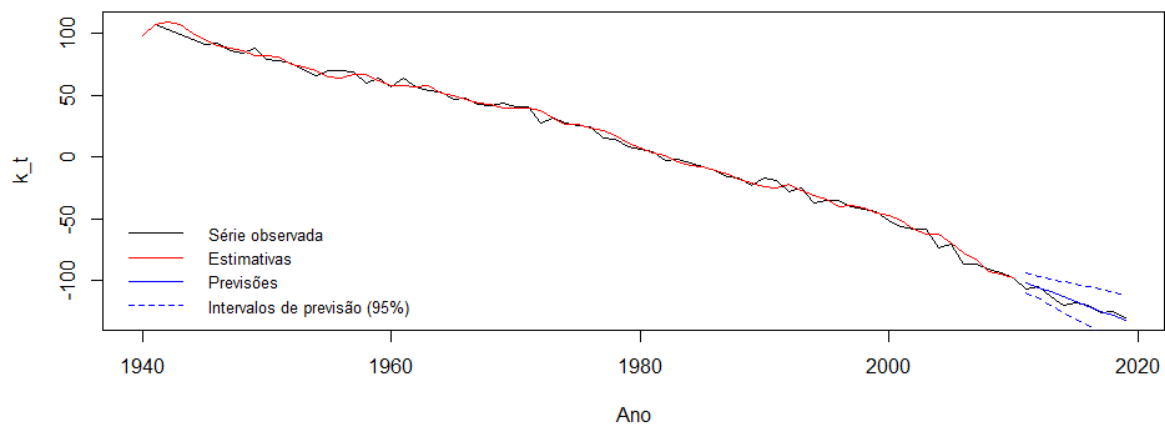


Figura 6.81: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sexo Feminino.

Por fim, no caso sem distinção de sexo, tal como nos anteriores, foram comparados os dois modelos obtidos através da aplicação do Método LC e a sua Extensão, sendo que os respetivos parâmetros de cada modelo e os valores de AIC e BIC estão apresentados na Tabela 6.26. Mediante esta verifica-se que o modelo obtido a partir do Método Lee-Carter apresenta o menor valor de AIC, contudo o modelo obtido através da Extensão do Método Lee-Carter exibe o menor valor de BIC. Uma vez que a partir desta

tabela não é retirada nenhuma conclusão, são calculadas as medidas de avaliação para a série temporal no período de treino e no período de teste e, consoante os resultados obtidos, verifica-se qual é o modelo mais adequado tanto para estimação como para previsão. Estes resultados estão apresentados na Tabela 6.27 e através desta constata-se que o Método Lee-Carter apresenta melhores resultados, ou seja, é o mais preciso tanto no período de treino como no período de teste, sendo por isso o modelo mais adequado dos dois apresentados.

Tabela 6.26: Modelos ARIMA mais adequados para cada método - Sem Distinção de Sexo.

Método	Modelo	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\phi}_4$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	AIC	BIC
Lee-Carter	ARIMA(2,2,3)	-1,63	-0,98	-	-	0,51	-0,60	-0,91	361,57	374,98
Extensão de LC	ARIMA(0,2,2)	-	-	-	-	-1,5	0,69	-	365,97	372,67

Tabela 6.27: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sem Distinção de Sexo.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	8,192	2,862	0,126	2,230	0,326	Série de Treino
Extensão de Lee-Carter	10,027	3,167	0,160	2,556	1,256	
Lee-Carter	14,457	3,802	0,039	3,333	1,266	Série de Teste
Extensão de Lee-Carter	92,248	9,605	0,079	8,042	2,291	

Contudo, como foi referido anteriormente, o melhor modelo para modelar os dados nem sempre é o mais adequado para realizar previsões e, por essa razão, neste caso também foram testados todos os modelos que apresentavam coeficientes estatisticamente significativos no caso da série de nível geral de mortalidade obtida através do Método Lee-Carter tal como na obtida através da Extensão do Método Lee-Carter, com o intuito de verificar qual é o mais adequado para realizar previsões. Para isso, como forma de comparação, foram calculadas as medidas de avaliação entre as previsões obtidas em cada modelo com as observações originais (Tabela 6.28).

Tabela 6.28: Medidas de avaliação calculadas para a série do nível geral de mortalidade no período de teste (2011-2019), com base nos resultados obtidos na aplicação da Modelação ARIMA - Sem Distinção de Sexo.

Modelo	EQM	REQM	EPAM	EAM	U-Theil	Método
ARIMA(0,2,0)	11,420	3,379	0,028	2,723	0,755	
ARIMA(1,2,0)	8,394	2,897	0,024	2,329	0,624	
ARIMA(0,2,1)	15,153	3,893	0,033	0,891	0,646	Extensão do Método LC
ARIMA(0,2,2)	92,248	9,605	0,079	8,042	2,291	
ARIMA(1,2,1)	10,628	3,260	0,027	2,639	0,713	
ARIMA(2,2,3)	14,457	3,802	0,039	3,333	1,266	Método LC

Na Tabela 6.28 verifica-se que os modelos aplicados à série do nível geral de mortalidade obtida a partir da aplicação da Extensão do Método LC, no geral, exibem uma qualidade preditiva melhor do que o modelo da série temporal obtida através do Método LC e, o modelo que apresenta a melhor qualidade preditiva é o modelo ARIMA(1,2,0), sendo ainda necessário testar os pressupostos dos resíduos do mesmo.

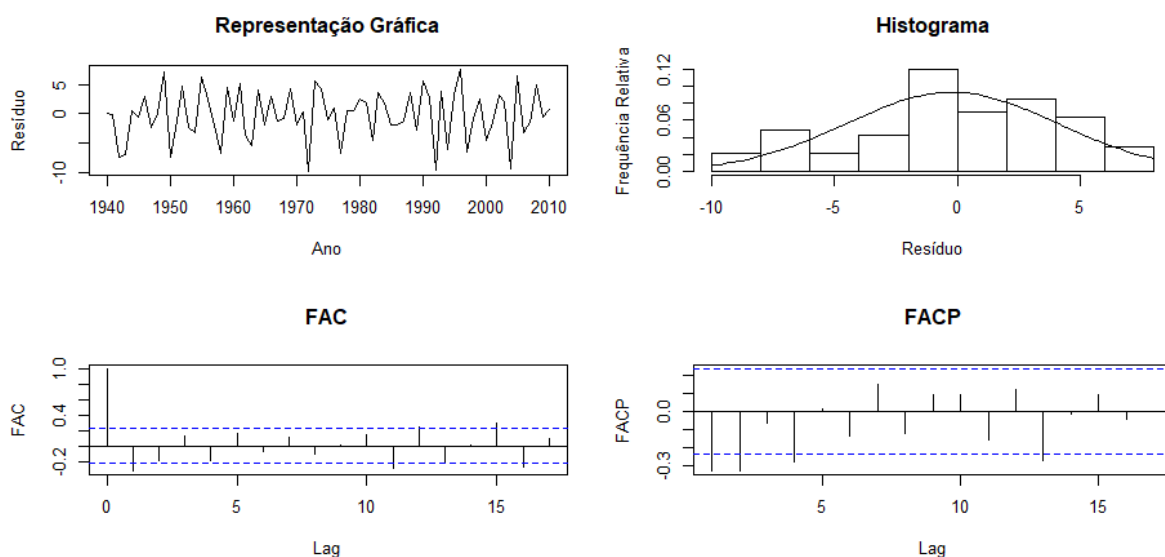


Figura 6.82: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.

Para o pressuposto da normalidade e do valor médio igual a zero foram utilizados os testes de *Lilliefors* e o teste *t*, respetivamente, sendo que no primeiro foi obtido um valor de prova igual a 0,505 e no

segundo igual a 0,685, ou seja, a hipótese de normalidade e de valor médio nulo não são rejeitadas. Já o pressuposto de independência entres os resíduos foi testado através do teste de *Ljung-Box*, obtendo-se um valor de prova igual a 0,013, ou seja, os resíduos são correlacionados. Por esta razão este modelo não é considerado adequado aos dados.

Posto isto, terão de ser testados os pressupostos dos resíduos do modelo alternativo, que neste caso é o modelo ARIMA(1,2,1). Na Figura 6.83 estão apresentados os gráficos necessários para a análise dos resíduos do modelo. Através do histograma da respetiva figura constata-se que os resíduos apresentam uma distribuição ligeiramente assimétrica à esquerda, porém através do teste de *Lilliefors* a hipótese de normalidade não é rejeitada (valor de prova igual a 0,487). Relativamente ao valor médio nulo, através do teste *t*, a hipótese nula não é rejeitada com um valor de prova igual a 0,068 e o pressuposto da variância constante pode ser observado através do gráfico dos resíduos que indica que estes oscilam uniformemente em torno de zero, presumindo que a variância é constante. O último pressuposto verificado é a existência de independência entre os resíduos com um valor de prova igual a 0,787 no teste de *Ljung-Box*. Portanto, depois de todos os pressupostos serem verificados, podemos concluir que o modelo é adequado aos dados e realizar as previsões através deste modelo. Estas estão apresentadas na Figura 6.84, onde se observa que as 9 observações que compõem a série de teste encontram-se inseridas no intervalo de previsão (taxa de cobertura igual a 100%).

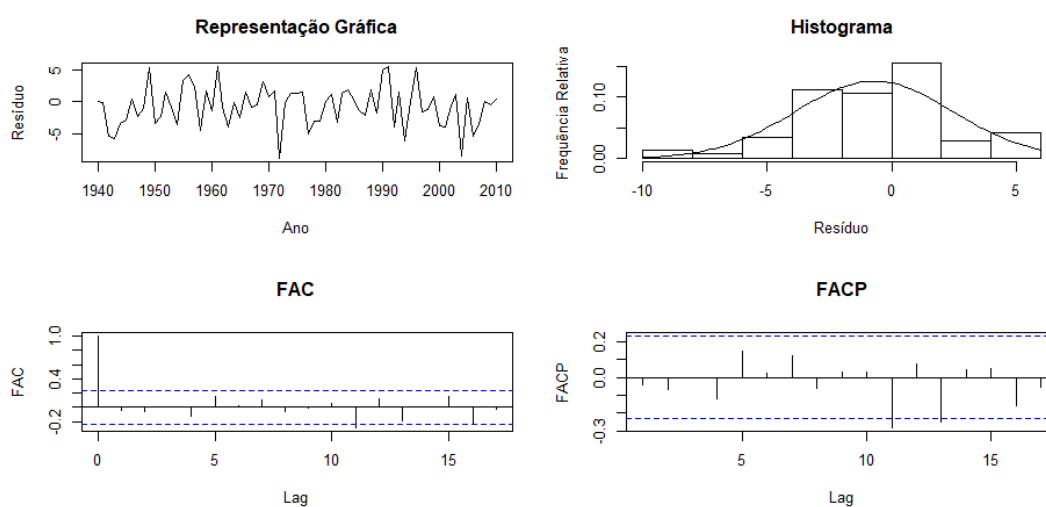


Figura 6.83: Série dos resíduos da série do nível geral de mortalidade após ajustamento do modelo ARIMA e respetivo histograma, FAC e FACP estimadas - Sem Distinção de Sexo.

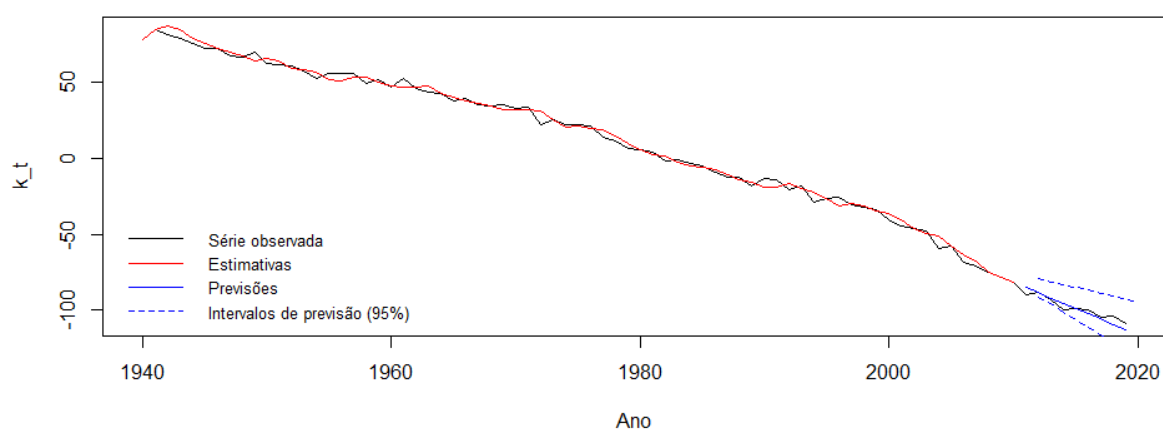


Figura 6.84: Previsões (no período de teste), pontuais e intervalares (95%), e estimativas pontuais (no período de treino) obtidas através do modelo ARIMA, sobrepostas à série do nível geral de mortalidade - Sem Distinção de Sexo.

Em suma, os modelos aplicados à série do nível geral de mortalidade obtida a partir da aplicação da Extensão do Método LC apresentam uma qualidade preditiva melhor nos três casos estudados, sendo o modelo mais adequado para realizar previsões o modelo ARIMA(1,2,1).

6.8.2 Método de Holt

Neste tipo de modelação os modelos obtidos são comparados também através das medidas de avaliação calculadas no período de treino e no período de teste, ou seja, através do EQM, REQM, EPAM, EAM e estatística de U-Theil.

Tabela 6.29: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Masculino.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	10,487	3,238	0,293	2,636	0,607	Série de Treino
Extensão de Lee-Carter	7,573	2,752	0,725	2,239	3,334	
Lee-Carter	73,374	8,507	0,081	6,686	1,520	Série de Teste
Extensão de Lee-Carter	96,111	9,804	0,089	8,224	2,510	

Na Tabela 6.29 estão apresentadas as medidas de avaliação calculadas no período de treino e no período de teste de ambas as séries e a partir destas constata-se que o modelo que melhor estima a série

do nível geral de mortalidade no período de treino (1940-2010) é o modelo da série temporal obtida pela aplicação da Extensão do Método Lee-Carter porém, para o período de teste (2011-2019), o modelo da série temporal obtida pela aplicação do Método Lee-Carter é o que apresenta melhores resultados.

Relativamente ao sexo feminino, as medidas de avaliação calculadas para ambos os períodos e para ambos os métodos são apresentadas na Tabela 6.30 e, mediante estas, observa-se que o modelo da série temporal obtida pelo Método Lee-Carter apresenta melhores resultados tanto na série de treino como na série de teste, comparativamente com o outro modelo.

Tabela 6.30: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Feminino.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	13,090	3,618	0,122	2,854	0,517	Série de Treino
Extensão de Lee-Carter	15,696	3,962	0,134	3,126	0,817	
Lee-Carter	39,187	6,260	0,063	5,858	1,214	Série de Teste
Extensão de Lee-Carter	119,475	10,930	0,075	9,116	2,197	

Por último, para o caso sem distinção de sexo, são apresentadas as cinco medidas de avaliação na Tabela 6.31 e através destas observa-se que, em relação à série de treino, a série temporal obtida através da Extensão do Método Lee-Carter apresenta melhores resultados em 3 das 5 medidas, já no período de teste o modelo que apresenta melhores resultados em todas as medidas de avaliação é o modelo da série temporal obtida a partir do Método Lee-Carter.

Tabela 6.31: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de treino (1940 a 2010) e no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação do Método de Holt - Sexo Feminino.

Método	EQM	REQM	EPAM	EAM	U-Theil	
Lee-Carter	10,732	3,276	0,144	2,611	0,484	Série de Treino
Extensão de Lee-Carter	10,014	3,164	0,162	2,557	1,309	
Lee-Carter	25,883	5,088	0,038	3,372	1,676	Série de Teste
Extensão de Lee-Carter	100,895	10,045	0,083	8,441	2,399	

Concluindo, em todos os casos seria melhor optar pelo método de Holt aplicado à série do nível geral de mortalidade obtida através do Método Lee-Carter para realizar previsões, pois este apresenta uma

melhor qualidade preditiva na série de teste.

6.8.3 Modelação ARIMA vs. Método de Holt

Depois de encontrados os modelos mais adequados aos dados utilizando os métodos de previsão paramétricos e não-paramétricos, falta somente averiguar qual dos dois é preferencial para realizar previsões para anos futuros. Para isso, são comparadas as medidas de avaliação obtidas em cada um nos três casos distintos. Na Tabela 6.32 encontram-se as medidas de avaliação dos modelos do sexo masculino, feminino e sem distinção de sexo. Através destas constata-se que os modelos produzidos pela modelação ARIMA são mais precisos do que os produzidos pelo método de Holt nos três casos estudados, uma vez que apresentam valores menores em todas as medidas de avaliação calculadas. Relativamente à taxa de cobertura dos intervalos de previsão, em todos os casos a taxa de cobertura é igual a 100%, ou seja, todas as observações da série original estão inseridas no respetivo intervalo de previsão.

Tabela 6.32: Medidas de avaliação calculadas para a série do nível geral de mortalidade, no período de teste (2011 a 2019), com base nos resultados obtidos na aplicação da Modelação ARIMA e Método de Holt.

Modelação	EQM	REQM	EPAM	EAM	U-Theil	Sexo
ARIMA	12,874	3,588	0,031	2,803	0,864	Masculino
Holt	72,374	8,507	0,081	6,686	1,520	
ARIMA	12,370	3,517	0,023	2,713	0,645	Feminino
Holt	39,187	6,260	0,063	5,858	1,214	
ARIMA	10,628	3,260	0,027	2,639	0,713	Sem Distinção
Holt	25,883	5,088	0,038	3,372	1,676	

7 Tábuas de Mortalidade da População

Portuguesa

Após modelado e previsto o nível geral de mortalidade torna-se possível calcular a taxa de mortalidade definida por $\ln(m_{x,t}) = \alpha_x + \beta_x k_t + \epsilon_{x,t}$ e, posteriormente, todas as funções básicas que integram uma tábua de mortalidade que são descritas no Capítulo 2.

Foram construídas as tábuas de mortalidade desde 1940 até 2019, com e sem distinção de sexo, elaboradas a partir do ajustamento da taxa de mortalidade proveniente da estimação dos parâmetros tanto do Método Lee-Carter como da Extensão do Método Lee-Carter (Modelo de Poisson), com o intuito de mais tarde serem comparadas com as tábuas de mortalidade realizadas com os dados da companhia. No Apêndice D, a título de exemplo, estão representadas as tábuas de mortalidade de 1970 para o sexo masculino, feminino e sem distinção de sexo, utilizando as duas metodologias para o ajustamento da taxa de mortalidade.

As funções mais relevantes para a empresa das tábuas de mortalidade são a taxa de mortalidade, $m_{x,t}$, e a esperança de vida, e_x . Nas Figuras 7.85, 7.86 e 7.87 está representada a evolução da taxa de mortalidade nos anos 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010 e 2019, para cada sexo e para os modelos aplicados a todas as idades. Observa-se que o comportamento da taxa de mortalidade é semelhante no caso do sexo masculino e feminino, isto é, da idade 0 até aos 5 anos, aproximadamente, apresenta um decréscimo, tornando-se quase constante até aos 55 anos e, de seguida, observa-se um aumento até aos 100 anos. Os valores mais elevados desta taxa em idades iniciais nos primeiros anos do estudo são explicados na Secção 6.1 e, tal como é referido nessa secção, a mortalidade infantil diminuiu ao longo dos anos. Por essa razão é que se observa uma diminuição nas taxas de mortalidade em idades iniciais ao longo dos anos de estudo. Constata-se também que existem oscilações menos acentuadas, ao longo das idades, na taxa de mortalidade ajustada pela Extensão do Método Lee-Carter e que, em geral, os valores da taxa de mortalidade são mais elevados no caso da utilização do Método Lee-Carter para o

ajustamento da mesma, apesar de não existir uma diferença muito notória entre os dois métodos.

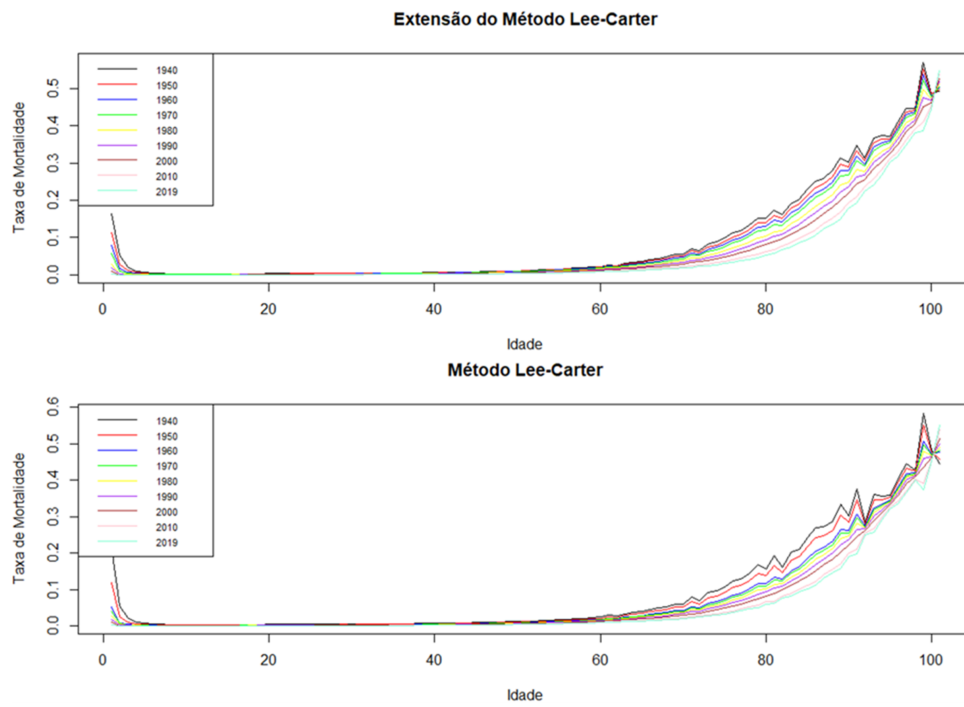


Figura 7.85: Evolução da taxa de mortalidade - Sexo Masculino.

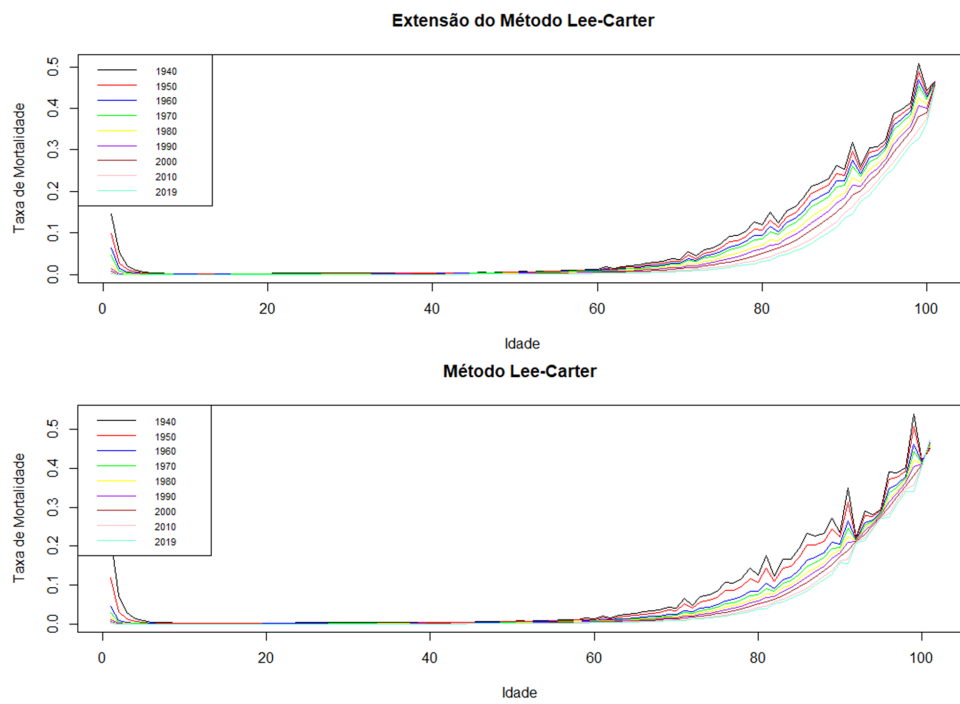


Figura 7.86: Evolução da taxa de mortalidade - Sexo Feminino.

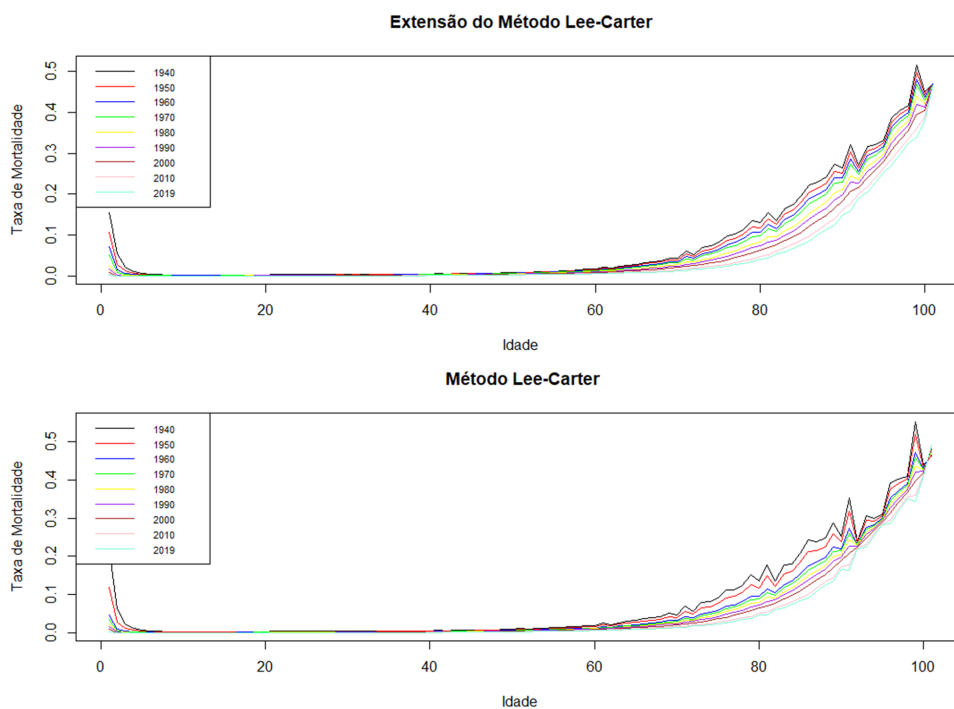


Figura 7.87: Evolução da taxa de mortalidade - Sem Distinção de Sexo.

Nas Figuras 7.88, 7.89 e 7.90 são apresentadas as curvas relativas à esperança de vida nos anos 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010 e 2019 para cada sexo, utilizando novamente os diferentes métodos de ajustamento da taxa de mortalidade. Após a observação das mesmas constata-se que a curva da esperança de vida ao longo do tempo apresenta um comportamento inverso ao da taxa de mortalidade, isto é, no caso da esperança de vida, esta diminui à medida que a idade aumenta, enquanto que a taxa de mortalidade aumenta com o aumento da idade, o que já era expectável. Nos primeiros anos da análise (1940 e 1950) observa-se que, para idades iniciais, a esperança de vida apresenta valores menores comparativamente com os últimos anos da análise. Isto verifica-se devido ao elevado número de mortes observadas na altura, como já foi referido anteriormente, sendo que, como ocorriam mais mortes nestas idades, a esperança de vida torna-se mais reduzida. Visto que, com o passar dos anos, a mortalidade infantil em Portugal diminuiu para valores bastante baixos, a esperança de vida em idades iniciais tornou-se mais elevada. Por exemplo, utilizando a Extensão do Método Lee-Carter, em 1940 a esperança de vida de uma pessoa acabada de nascer seria de 52 anos, enquanto que uma pessoa que nasceu em 2019 teria uma esperança de vida de, aproximadamente, 82 anos. Posto isto, pode-se afirmar que a esperança de vida é inversamente proporcional ao número de mortes, isto é, à medida que o número de mortes aumenta, a esperança de vida diminui e vice-versa. É possível também verificar que

ao longo dos anos a esperança de vida tem vindo a aumentar para todas as idades, sendo que só a partir dos 90 anos é que a mesma é muito semelhante em todos os anos do estudo.

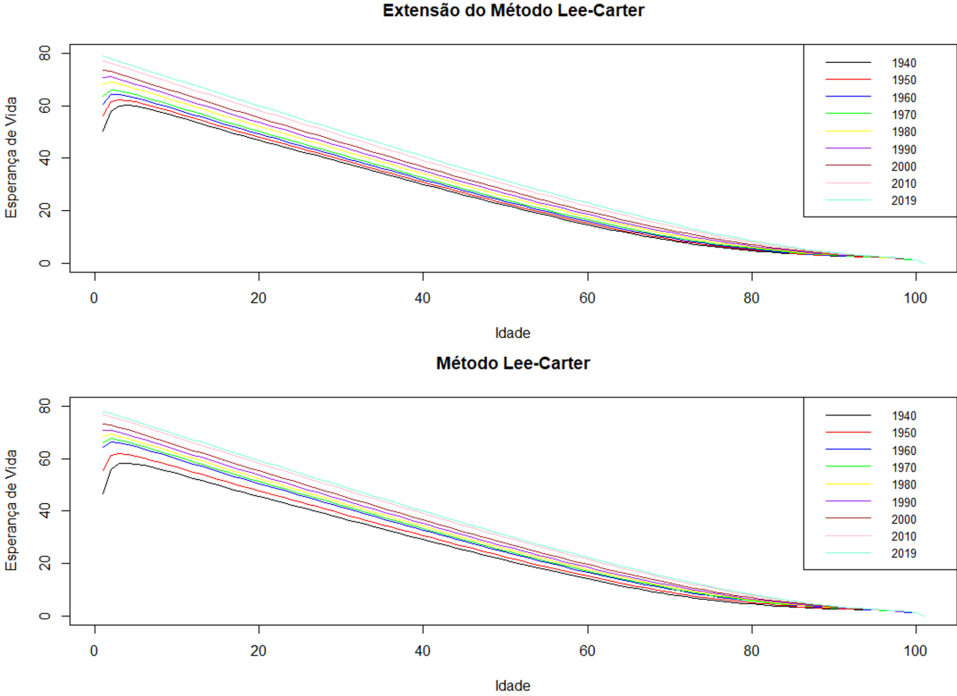


Figura 7.88: Evolução da esperança de vida - Sexo Masculino.

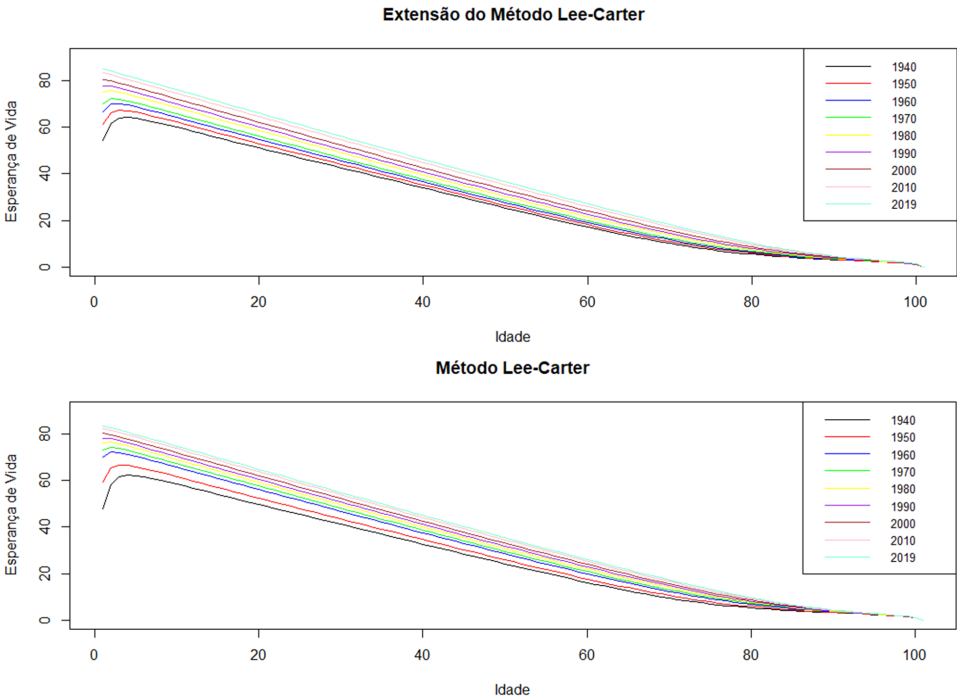


Figura 7.89: Evolução da esperança de vida - Sexo Feminino.

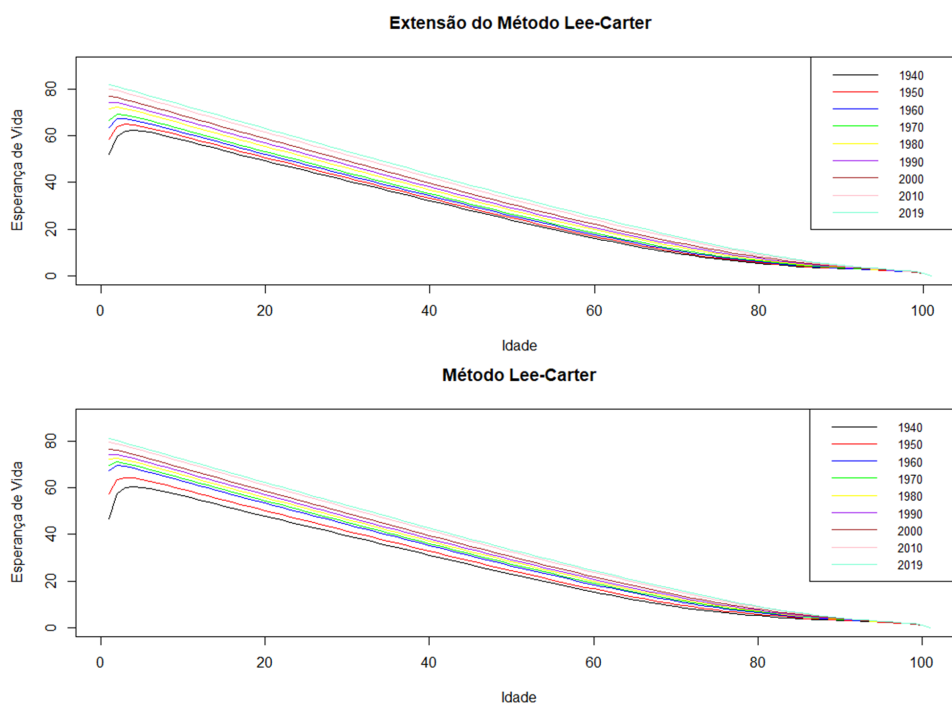


Figura 7.90: Evolução da esperança de vida - Sem Distinção de Sexo.

Antes de se efetuar a comparação dos métodos de ajustamento utilizados, é necessário referir que o objetivo da aplicação destes métodos é encontrar o método que produz os menores valores de esperança de vida, para que as taxas vitalícias, que irão ser abordadas na Secção 7.1, apresentem valores mais reduzidos. Ao realizar a comparação dos resultados obtidos entre os dois métodos utilizados, constata-se que a esperança de vida calculada através da taxa de mortalidade ajustada pelo Método Lee-Carter é menor comparativamente com a calculada através do outro método, contudo é de realçar que as diferenças existentes entre as esperanças de vida obtidas pelos dois métodos não são muito expressivas.

7.1 Cálculo das Pensões Vitalícias

Um dos objetivos deste trabalho é desenvolver modelos que permitam estimar os fatores necessários para o cálculo das pensões vitalícias do ramo de acidentes de trabalho. Este tipo de pensões são aplicadas a clientes que apresentem uma redução na capacidade de trabalho ou de ganho do trabalhador e são pagas consoante o grau de incapacidade do mesmo.

O cálculo destas pensões é obtido pela fórmula de matemática financeira apresentada em 7.1, onde e_x representa a esperança de vida, TE a taxa de encargos e T_x a taxa vitalícia. A taxa vitalícia consiste no fator que irá ser aplicado à pensão anual do sinistrado, isto é, ao valor anual que o mesmo recebe pela

incapacidade que sofreu, de modo a calcular a pensão vitalícia do mesmo. Já a taxa de encargos consiste numa percentagem que a própria companhia aplica sobre as provisões matemáticas para fazer face aos custos que a mesma dispõe com a gestão das pensões. Por uma questão de simplificação, neste estudo, esta será assumida como constante e igual a 0% para todos os cálculos e, assim a taxa vitalícia é dada por

$$T_x = \left(e_x + \frac{13}{24} \right) \times (1 + TE). \quad (7.1)$$

Neste trabalho foram calculadas as taxas vitalícias para cada tipo de sexo desde 1940 até 2019, utilizando cada um dos métodos de modelação da taxa de mortalidade referidos nas Secções 6.2 e 6.5. A título de exemplo, no Apêndice E estão apresentadas as tabelas referentes às taxas vitalícias calculadas para cada tipo de sexo no ano 1970. Uma vez que o cálculo da taxa depende do valor da esperança de vida, é expectável que os métodos de modelação que apresentaram esperanças de vida menores apresentem também taxas com valores menores, comparativamente com as restantes.

De forma a entender os valores obtidos e identificados nas tabelas do Apêndice E, consideremos uma situação hipotética. Um homem que tenha nascido em 1970 e que apresente uma pensão anual igual a 12 000 €, caso tenha sofrido um acidente de trabalho, será criada uma provisão de 268 272 € para o mesmo, no caso de ser utilizada a tábua de mortalidade de 1970 calculada através da taxa de mortalidade obtida pela Extensão do Método Lee-Carter, ou de 276 156 €, no caso de ser utilizada a tábua de mortalidade de 1970 calculada através da taxa de mortalidade obtida pelo Método Lee-Carter. Isto significa que se fosse utilizada a tábua de mortalidade obtida através da taxa de mortalidade ajustada pelo Método Lee-Carter, ao invés da obtida a partir da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter, observava-se um impacto de 2,9%.

Tabela 7.33: Cálculo de uma pensão vitalícia.

Método	Taxa Vitalícia (T_x)	Pensão Anual (P)	Pensão Vitalícia ($T_x \times P$)
Lee-Carter	23,013	12 000 €	276 156 €
Extensão de LC	22,356	12 000 €	268 272 €

8 Aplicação aos Dados da Seguradora

Para além de observar o comportamento da população portuguesa, é também necessário entender o comportamento da população da companhia, para entender se o comportamento é semelhante nas duas populações. Para tal, neste capítulo é realizada inicialmente uma análise exploratória dos dados da companhia e, posteriormente, são calculadas as taxas de mortalidade para que possam ser comparadas com as obtidas anteriormente com os dados da população portuguesa.

8.1 Análise Descritiva dos Dados

Nesta secção são apresentados os dados da seguradora utilizados, relativos ao número de mortes e à exposição ao risco nos anos 2007-2021. Tal como nos dados da população portuguesa, é feita a análise dos dados com e sem distinção de sexo através de gráficos de linhas, para averiguar o comportamento de cada série temporal, e de seguida são elaborados *boxplots* para observar a distribuição dos dados e se estes apresentam valores discrepantes.

8.1.1 Caso 1: Sexo Masculino

Relativamente ao número de pessoas expostas ao risco em cada ano, através da Figura 8.91 observa-se que a curva apresenta uma tendência muito semelhante ao longo dos anos, sendo que em todos os anos este valor atinge o valor máximo na faixa etária dos 40-65 anos. A única diferença é que este valor máximo tende, com o passar dos anos, a ser atingido em idades cada vez maiores, por exemplo, em 2007 a idade que apresentava uma maior exposição ao risco era 49 anos, em 2014 era 54 anos e, por fim, em 2021 era 62 anos.

Na Figura 8.92 são apresentados os *boxplots* de cada ano e a partir destes constata-se que os referentes aos anos 2020 e 2021 apresentam um *outlier* cada. Porém, depois de analisados,

reconhece-se que estes não são considerados erros mas sim valores extremos, uma vez que ambos representam o valor máximo de cada conjunto de observações.

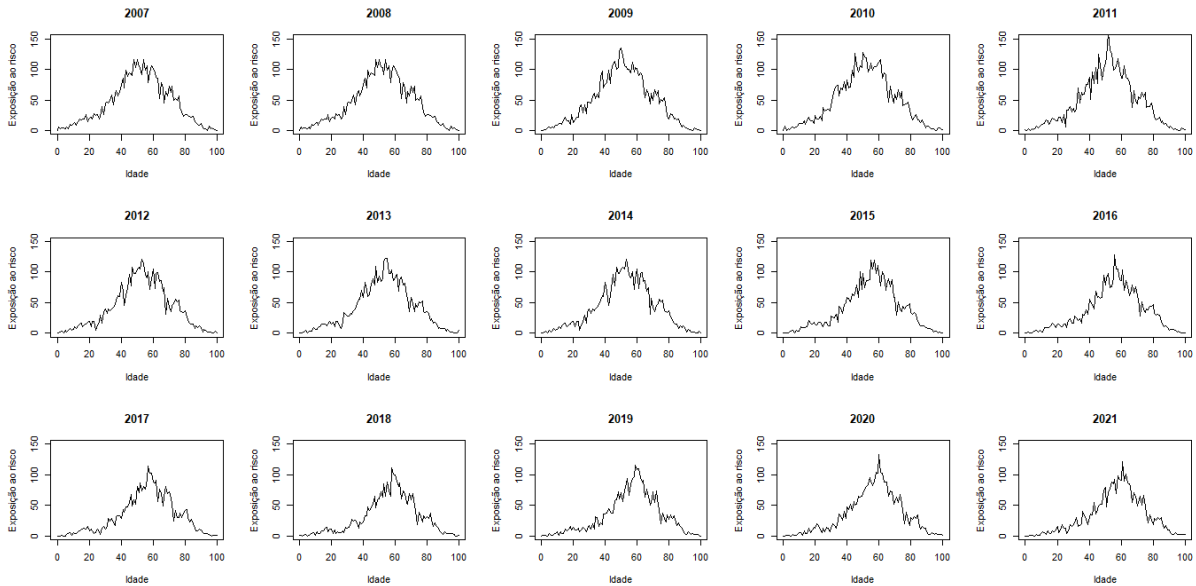


Figura 8.91: Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Masculino.

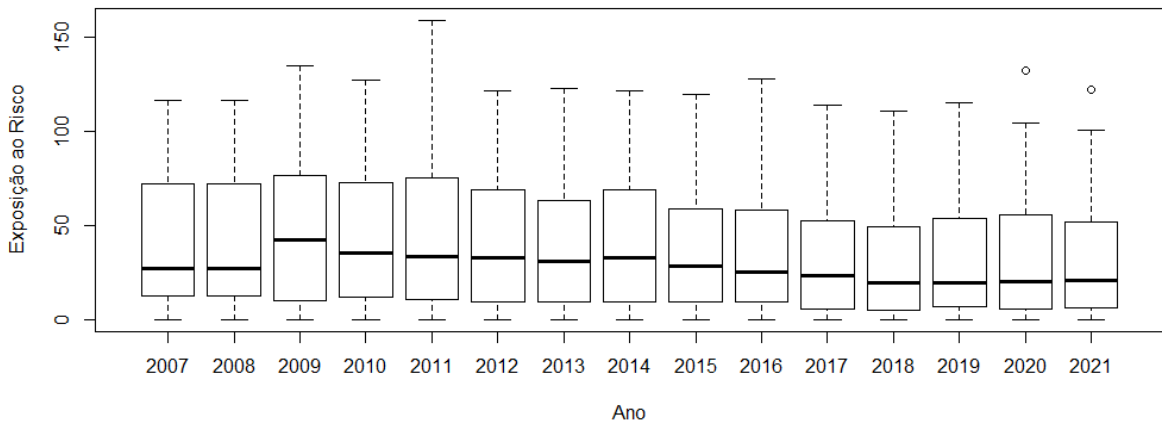


Figura 8.92: *Boxplots* do Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Masculino.

No que diz respeito ao número de mortes, entre os anos 2009 e 2021 o comportamento ao longo da idade é muito semelhante, sendo que o pico máximo apresenta-se na faixa etária dos 65-90 anos, com a exceção do ano 2016 em que o valor máximo é atingido aos 60 anos. Os anos 2007 e 2008 foram os anos onde se observaram menos números de mortes.

Com o intuito de averiguar a existência de *outliers* nos dados, foram elaborados os *boxplots* apresentados na Figura 8.94. Mediante esta observa-se que a maior parte dos dados apresentam valores discrepantes, contudo estes valores não são considerados erros, são considerados valores extremos. Isto acontece devido à existência de vários valores nulos em várias idades, o que faz com que estes valores sejam considerados como extremos.

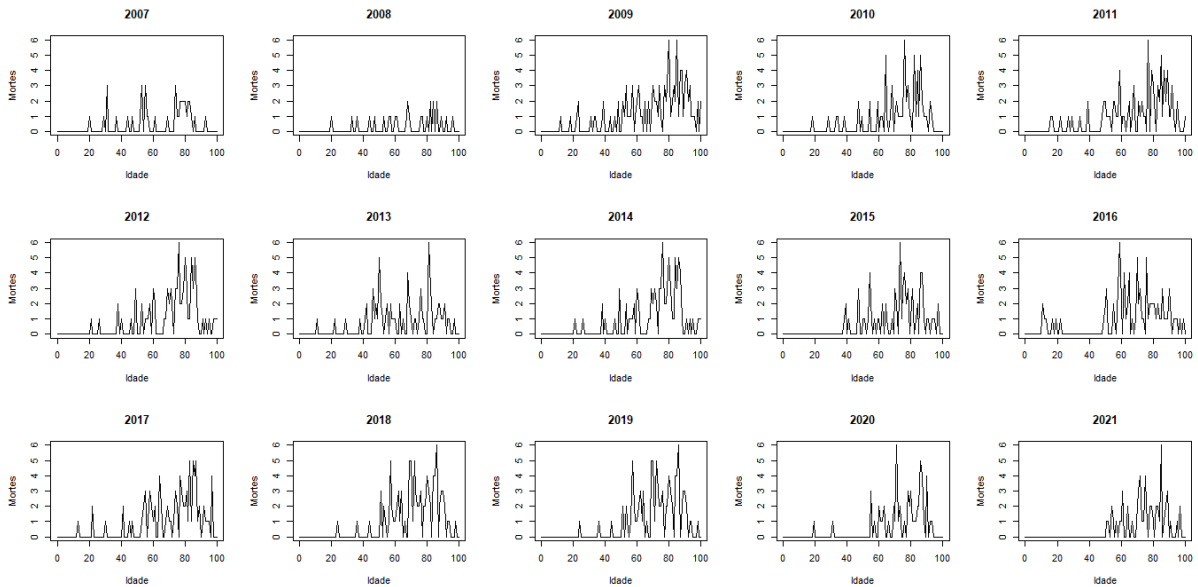


Figura 8.93: Número de Mortes nos Anos 2007-2021 - Sexo Masculino.

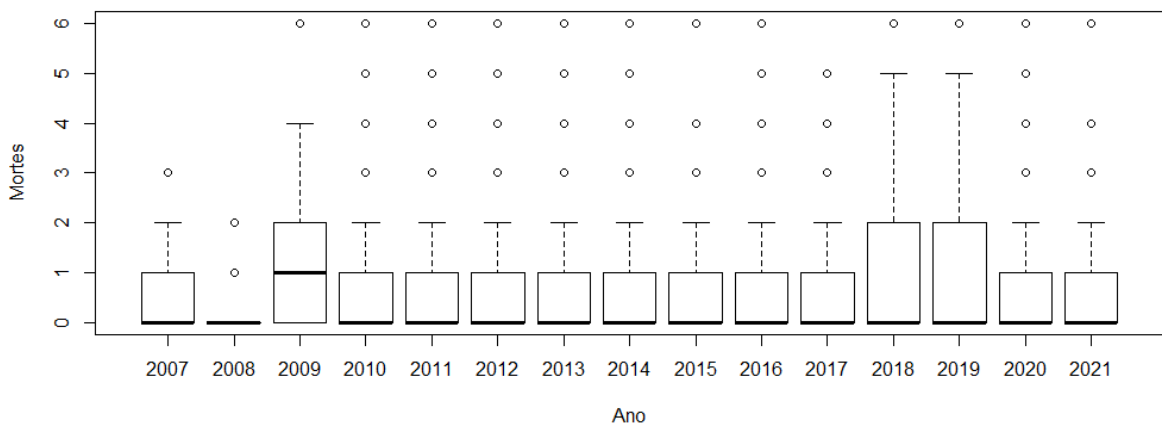


Figura 8.94: *Boxplots* do Número de Mortes nos Anos 2007-2021 - Sexo Masculino.

8.1.2 Caso 2: Sexo Feminino

Tal como no primeiro caso, a curva do número de expostos ao risco no caso do sexo feminino apresenta um comportamento semelhante ao longo dos anos, como se pode observar na Figura 8.95, alterando apenas o valor máximo em cada ano. Através da figura constata-se que a idade onde existe um maior valor desta variável vai aumentando com o tempo, sendo que no ano 2007 o maior valor era observado em mulheres com 57 anos, em 2014 com 61 anos e em 2021 com 64 anos.

No que respeita à existência de *outliers*, através da Figura 8.96 observa-se apenas a existência destes nos anos 2016 e 2021 e, após a análise dos mesmos, verifica-se que estes são apenas valores extremos que correspondem ao máximo dos respetivos anos. Mediante a mesma figura também é possível averiguar a distribuição dos dados e percebe-se que o ano que apresenta uma maior dispersão dos dados é o ano 2016 e constata-se também que, em todos os anos, as observações apresentam uma distribuição assimétrica à direita, ou seja, uma assimetria positiva que ocorre quando o valor da mediana é mais próximo do valor do 1º quartil.

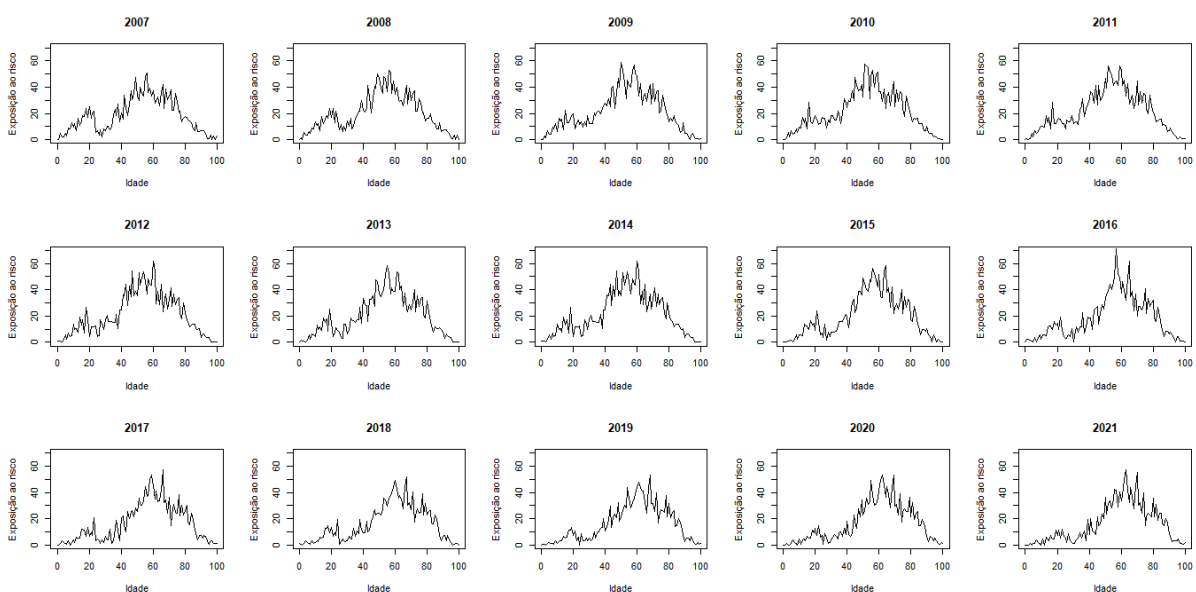


Figura 8.95: Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Feminino.

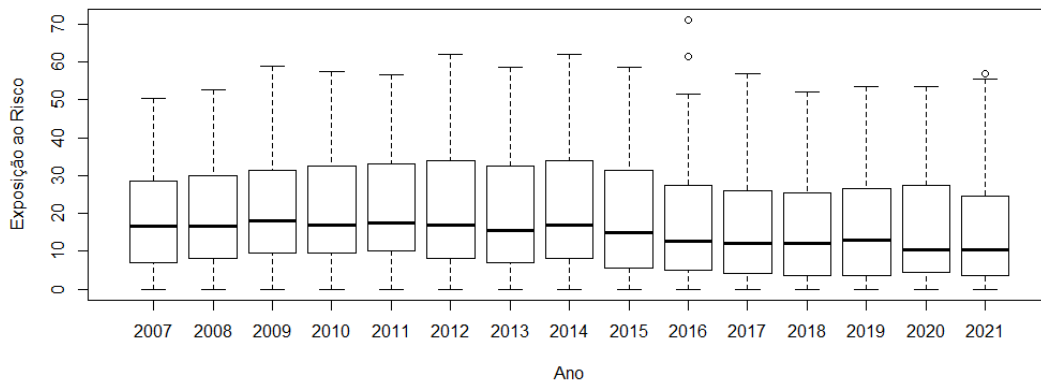


Figura 8.96: *Boxplots* do Número de Expostos ao Risco nos Anos 2007-2021 - Sexo Feminino.

Relativamente ao número de mortes observa-se que este valor é inferior na maior parte dos anos, comparativamente com os valores do sexo masculino, e que, em geral, a faixa etária onde se verifica um maior número de mortes é entre os 70 e 100 anos. O elevado número de *outliers* visíveis nos *boxplots* da Figura 8.94 deve-se à existência de diversos valores nulos na maior parte das idades, o que faz com que nas idades onde se observam valores positivos essas observações sejam consideradas como valores extremos.

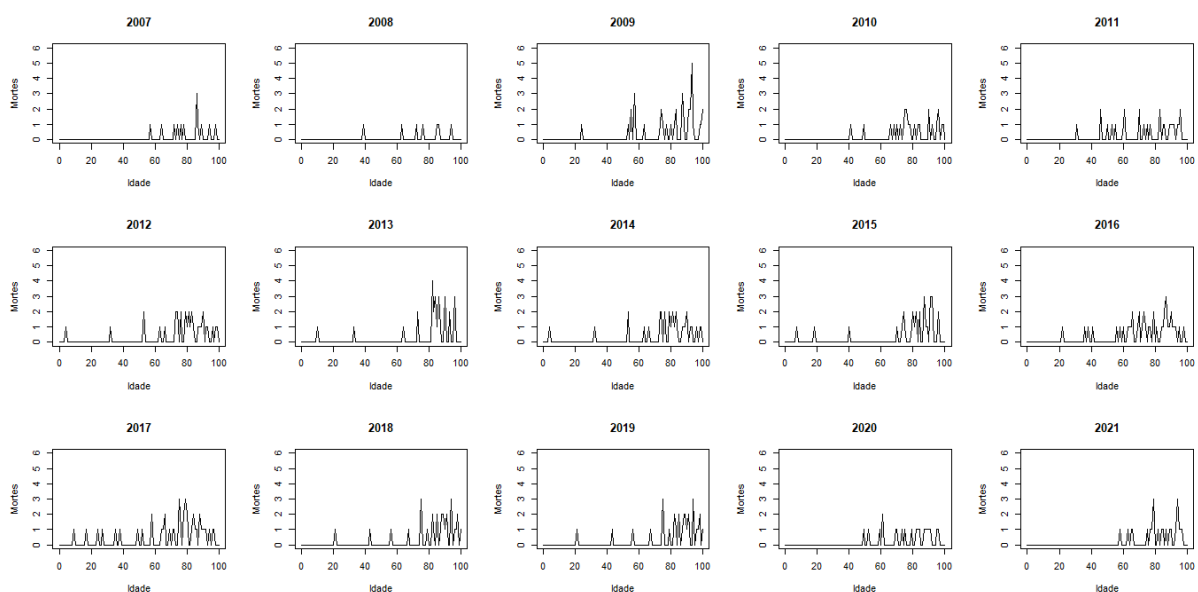


Figura 8.97: Número de Mortes nos Anos 2007-2021 - Sexo Feminino.

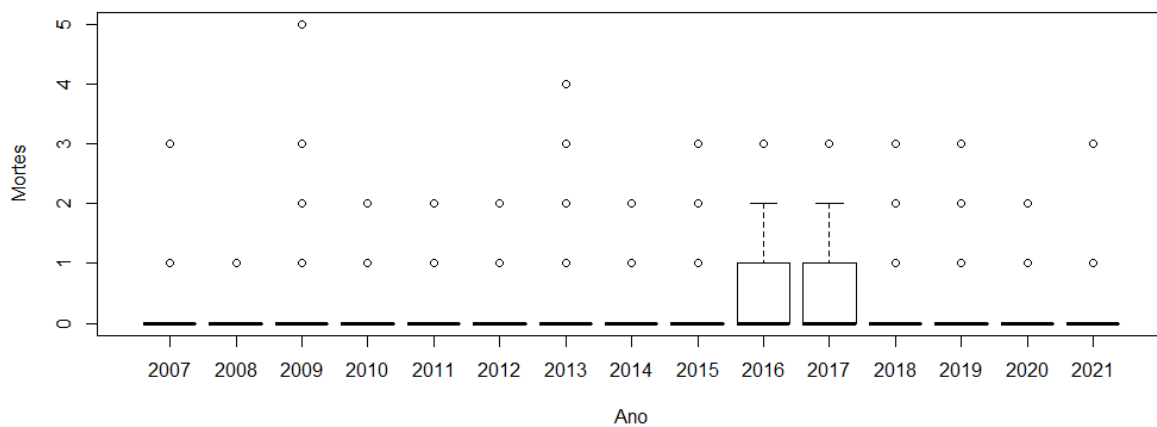


Figura 8.98: *Boxplots* do Número de Mortes nos Anos 2007-2021 - Sexo Feminino.

8.1.3 Caso 3: Sem Distinção de Sexo

O terceiro e último caso é quando não é feita a distinção entre sexos, ou seja, é o total das observações mencionadas anteriormente. Neste caso, e em relação ao número de expostos ao risco, na Figura 8.99 é possível observar o comportamento desta variável ao longo do tempo e da idade, onde se constata que o comportamento da mesma é semelhante ao longo dos anos, isto é, apresenta valores menores entre as faixas etárias dos 0-40 anos e dos 65-100 anos e entre a faixa etária dos 40-65 anos é onde atinge o máximo, sendo que este varia ao longo do tempo.

Quanto à existência de *outliers*, neste caso não existem valores extremos. Na Figura 8.100 também é possível verificar como se distribuem os dados ao longo dos anos e observa-se que o ano que apresenta uma maior dispersão é o ano 2011 e que as observações de todos os anos apresentam uma distribuição assimétrica à direita, ou seja, uma assimetria positiva.

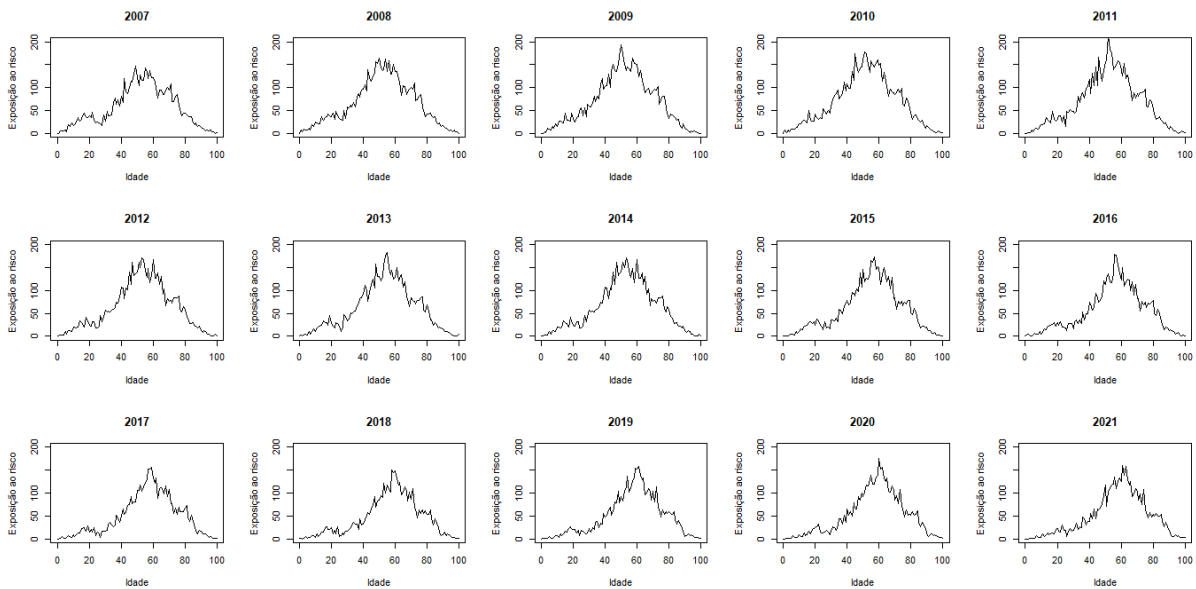


Figura 8.99: Número de Expostos ao Risco nos Anos 2007-2021 - Sem Distinção de Sexo.

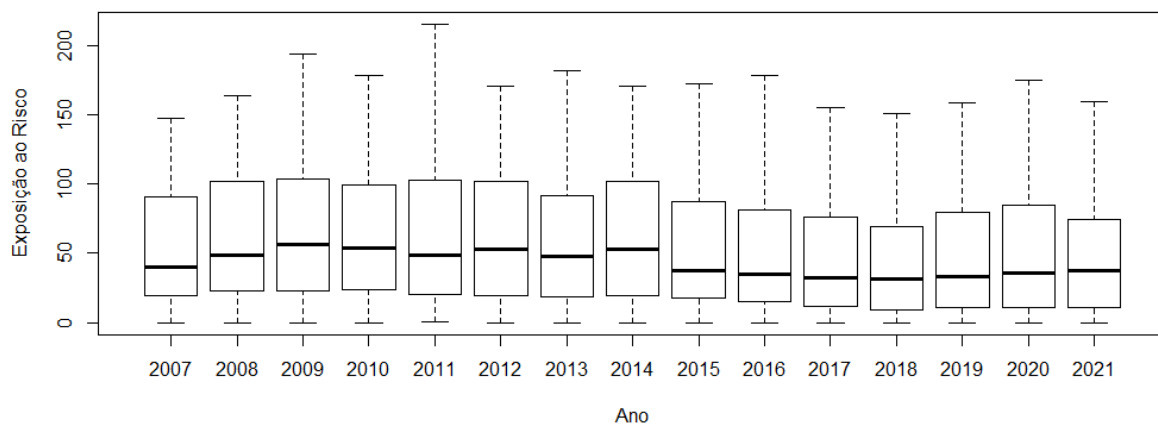


Figura 8.100: *Boxplots* do Número de Expostos ao Risco nos Anos 2007-2021 - Sem Distinção de Sexo.

Em relação ao número de mortes e através da Figura 8.101, percebe-se que não foram verificadas tantas mortes nos anos 2007 e 2008, comparativamente com os restantes anos, e que o valor máximo atingido varia muito de ano para ano, sendo que a faixa etária que compreende todos estes valores é a dos 70-95 anos.

Na Figura 8.102 observa-se que existe um grande número de *outliers* em todos os anos contudo, como já foi referido anteriormente, isto acontece devido ao elevado número de observações nulas que existem em cada ano, fazendo com que os valores mais elevados sejam considerados como extremos e

não como erros. O elevado número de observações nulas faz também com que a distribuição dos dados em muitos anos seja assimétrica positiva e com mediana igual a zero, como acontece nos anos 2007, 2008, 2010, 2012, 2013, 2018, 2019, 2020 e 2021.

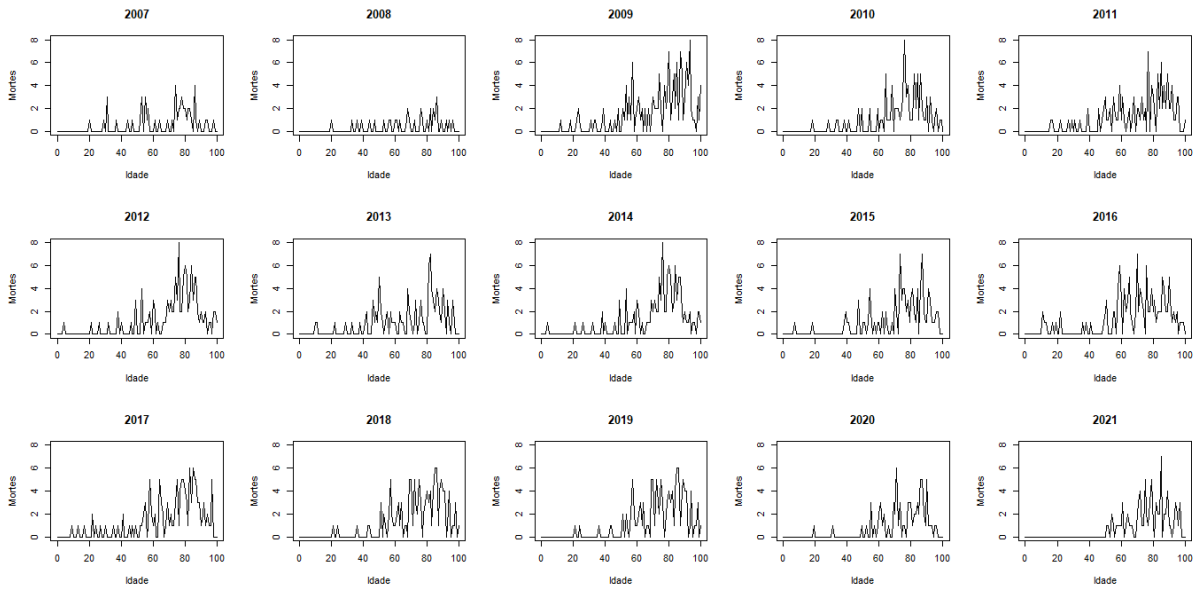


Figura 8.101: Número de Mortes nos Anos 2007-2021 - Sem Distinção de Sexo.

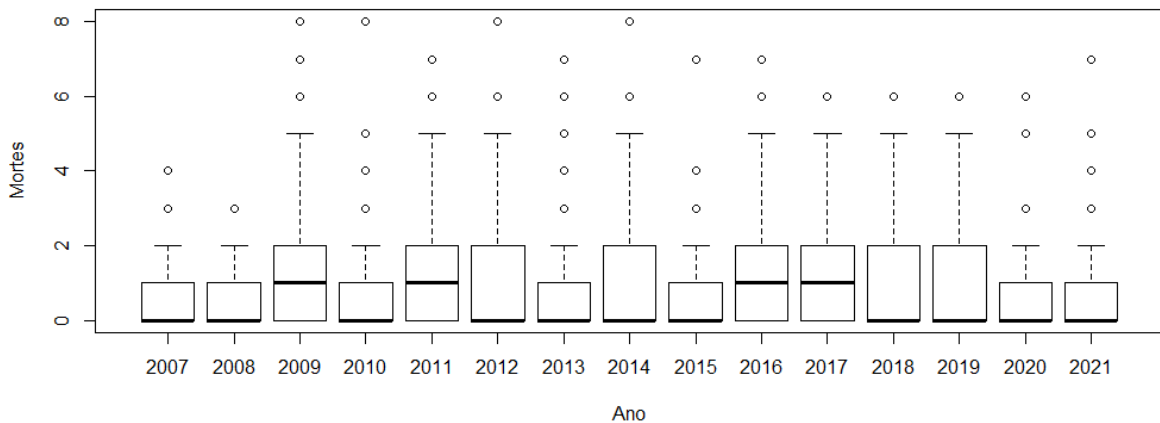


Figura 8.102: *Boxplots* do Número de Mortes nos Anos 2007-2021 - Sem Distinção de Sexo.

Por fim, de modo a completar a análise gráfica apresentada anteriormente, foram calculadas algumas medidas descritivas aos dados da exposição ao risco e do número de mortes de cada sexo (Tabelas F.1 e F.2 do Apêndice A). Com o mesmo intuito foram também calculadas as médias anuais de cada variável

para averiguar o comportamento das mesmas ao longo do tempo e estas estão disponíveis nas Tabelas F.3 e F.4 do Apêndice A.

Relativamente à exposição ao risco, através da Tabela F.1 do Apêndice A observa-se que o sexo masculino apresenta um valor máximo superior e que este ultrapassa o dobro do valor máximo do sexo feminino. Além disso, em média, o número de expostos ao risco do sexo masculino é superior ao do sexo feminino ao longo do período de estudo e observa-se que, no sexo masculino, os dados apresentam uma tendência crescente entre 2007-2010 e 2019-2021 e uma tendência decrescente entre 2011-2018, já no sexo feminino verifica-se uma tendência crescente entre 2007-2011 e nos restantes anos observa-se um comportamento irregular, ou seja, observam-se decréscimos seguidos de aumentos e vice versa. Para verificar a dispersão dos dados, através dos valores da variância constata-se que os dados do sexo masculino apresentam uma variabilidade muito superior à dos dados do sexo feminino.

No que respeita ao número de mortes, como foi referido anteriormente, verifica-se a existência de valores nulos em muitas idades, sendo que, no caso do sexo feminino, pelo menos 75% das observações correspondem a valores nulos, já no caso do sexo masculino, 75% das observações estão situadas abaixo de 1. Observando as médias das observações de cada ano, verifica-se em ambos os sexos um comportamento irregular, ou seja, não apresentam uma tendência linear ao longo dos anos, porém, em média, o número de mortes do sexo masculino é superior ao número do sexo feminino. Por último, verifica-se que os dados do sexo masculino apresentam uma variabilidade maior do que os dados do sexo feminino.

8.2 Tratamento de Dados - Médias Móveis e Médias Móveis Ponderadas

Como foi referido na Secção 8.1, um dos problemas dos dados da companhia é a existência de valores nulos em muitas idades relativas a vários anos. Para ultrapassar este problema, foram propostas duas abordagens: a utilização de Médias Móveis e Médias Móveis Ponderadas. A única diferença entre as duas metodologias é que na utilização das Médias Móveis é atribuído o mesmo peso para todas as observações, enquanto que na utilização das Médias Móveis Ponderadas é atribuído um peso diferente consoante a aproximação ao valor que se pretende estimar. Para ser mais fácil de entender, na Tabela 8.34 estão apresentadas as equações utilizadas em cada um dos métodos, onde X_i representa a observação correspondente ao ano i ou idade i . Neste caso, inicialmente, estes métodos foram utilizados com um período de 5 anos e, posteriormente, com um período de 5 idades, tanto nos dados relativos à exposição ao risco como ao número de mortes.

Tabela 8.34: Equações das Médias Móveis e Médias Móveis Ponderadas.

Médias Móveis	$X'_i = (X_{i-2} + X_{i-1} + X_i + X_{i+1} + X_{i+2})/5$
Médias Móveis Ponderadas	$X'_i = (1X_{i-2} + 2X_{i-1} + 3X_i + 2X_{i+1} + 1X_{i+2})/9$

Depois de transformados os dados originais através das duas metodologias acima mencionadas, foram calculadas as taxas de mortalidade através da equação $d_{x,t}/E_{x,t}$. As taxas de mortalidade calculadas através das Médias Móveis e Médias Móveis Ponderadas, para cada ano e para cada sexo, estão apresentadas nas Figuras 8.103, 8.104 e 8.105. Através destas constata-se que nos três casos o comportamento da taxa de mortalidade é semelhante, sendo que esta apresenta menores valores para idades entre 0 e 70 anos e valores maiores para idades superiores a 70 anos. Comparando o sexo masculino com o sexo feminino, observa-se que, no geral, o sexo feminino apresenta uma taxa de mortalidade superior ao longo dos anos. Relativamente às metodologias utilizadas para combater o problema dos valores nulos existentes em alguns anos, não se observam grandes diferenças nos valores das taxas de mortalidade calculadas em cada um dos casos.

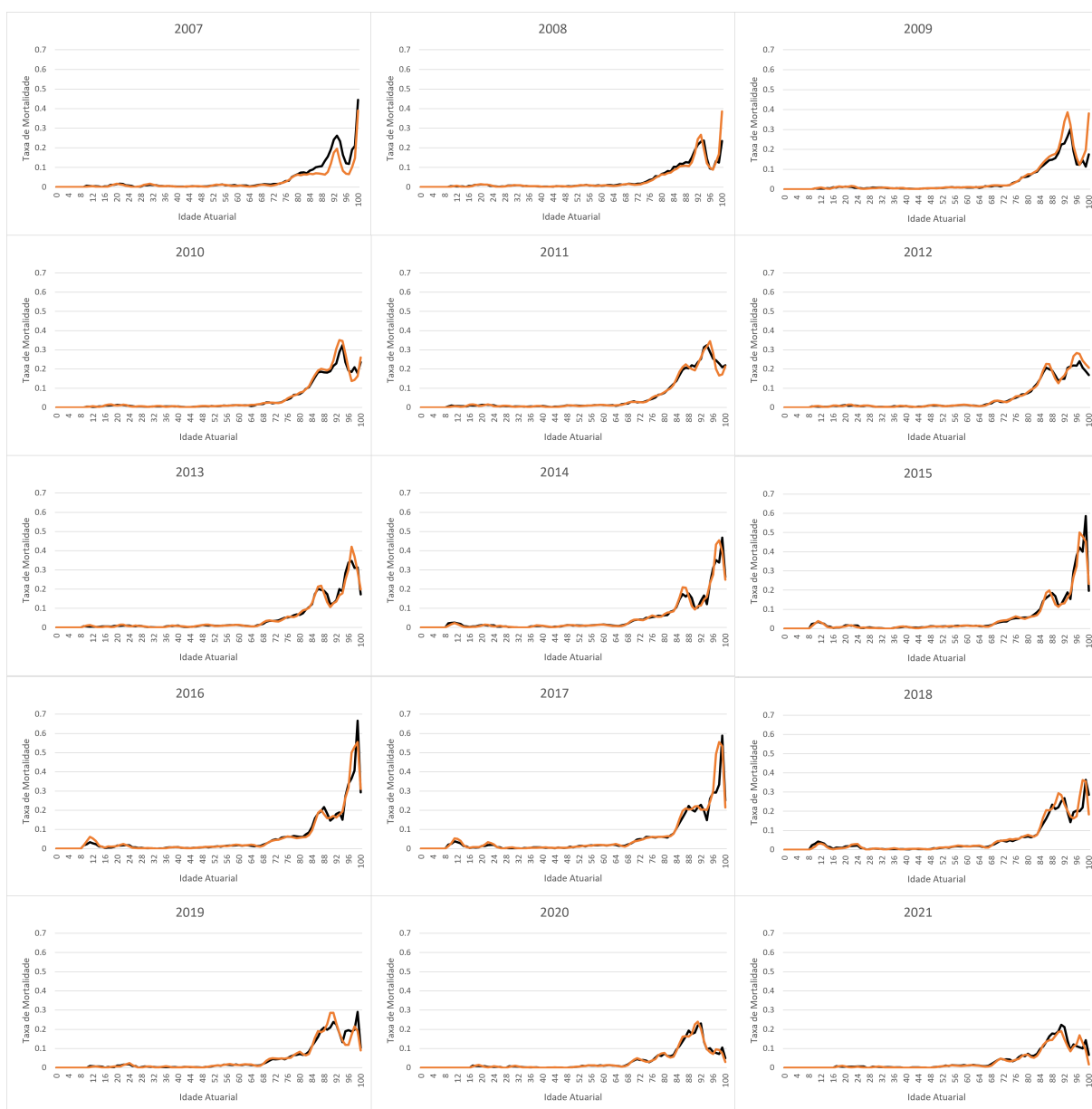


Figura 8.103: Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sexo Masculino.

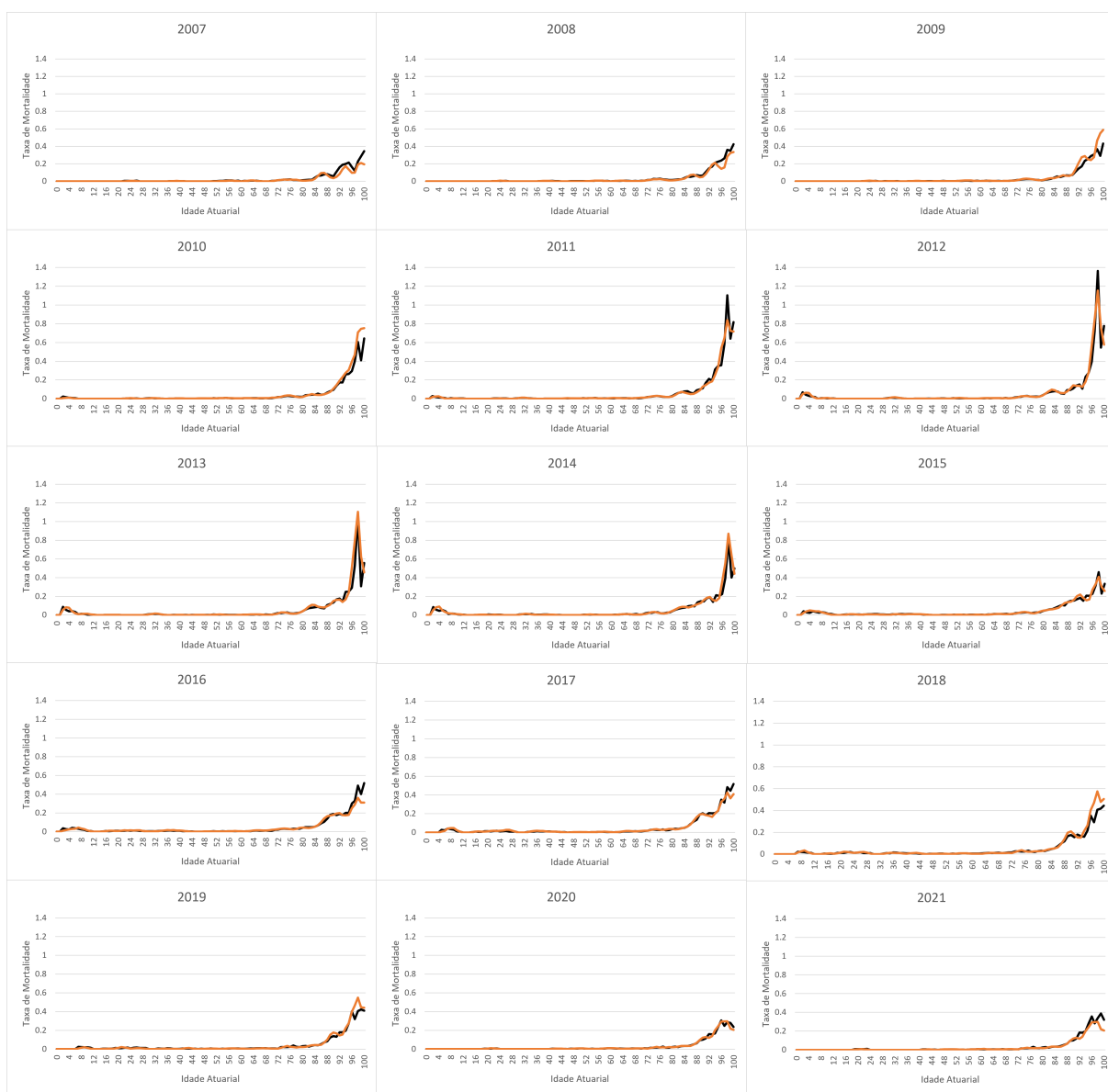


Figura 8.104: Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sexo Feminino.

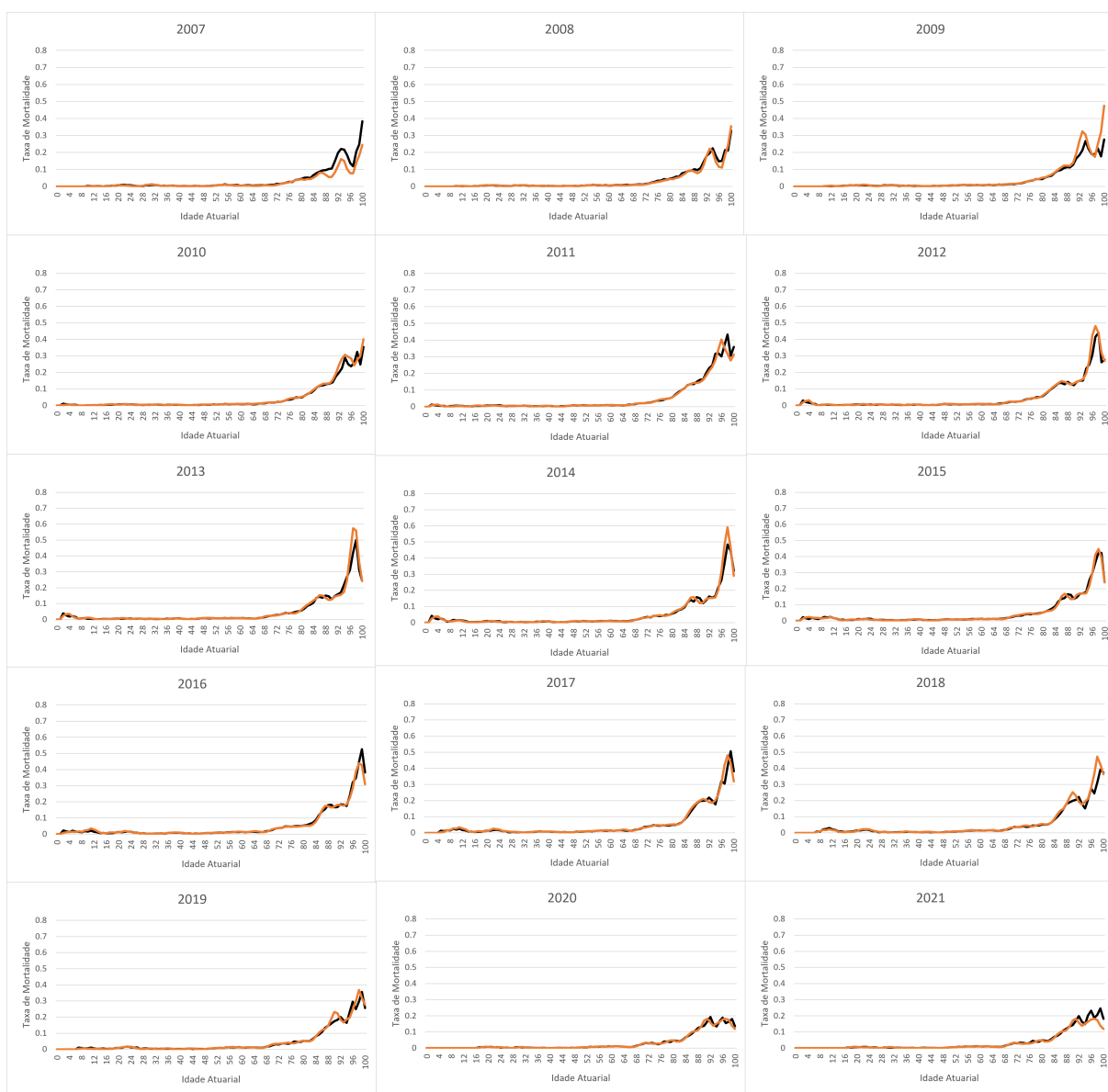


Figura 8.105: Evolução da taxa de mortalidade calculada através das Médias Móveis (linha preta) e Médias Móveis Ponderadas (linha laranja) - Sem Distinção de Sexo.

De forma a analisar se o comportamento dos pensionistas é semelhante ao comportamento da população portuguesa, é necessário comparar as taxas de mortalidade obtidas em cada um dos casos através do cálculo de medidas de avaliação. Uma vez que os dados da companhia apresentam um número de observações bem mais reduzido do que os dados da população portuguesa, para realizar a comparação com os dados da população portuguesa, decidiu-se utilizar como dados da companhia as médias das observações dos anos 2015, 2016, 2017 e 2018, tanto no caso dos dados obtidos por Médias Móveis como a partir de Médias Móveis Ponderadas. Depois de tratados os dados, a título de

exemplo, estes foram comparados com as taxas de mortalidade do ano 1997 da população portuguesa.

Nas Figuras 8.106 e 8.108 apresentam-se as curvas das taxas de mortalidade de cada ano e de cada população para todas as idades, já nas Figuras 8.107 e 8.109 estão apresentadas as mesmas curvas porém somente dos 25 aos 65 anos, sendo que esta é a faixa etária que contém mais pensionistas no ramo de acidentes de trabalho. Fazendo a comparação entre a curva relativa à taxa de mortalidade da população da companhia com as curvas relativas às taxas de mortalidade da população portuguesa, observa-se que o comportamento das curvas em ambos os sexos é semelhante até aos 75 anos e a partir desta idade é onde se observam mais diferenças. Contudo, comparando o comportamento das curvas nas Figuras 8.107 e 8.109, constata-se que a curva relativa ao sexo masculino apresenta um comportamento mais idêntico à curva relativa à população portuguesa, comparativamente com a curva do sexo feminino.

Isto pode ser explicado devido à existência de mais dados da companhia no caso do sexo masculino comparativamente com o sexo feminino. Na Figura 8.110 observa-se a comparação entre as frequências relativas dos dados da exposição ao risco existentes da população portuguesa e da população da companhia em ambos os sexos. Através desta constata-se que, relativamente aos dados da companhia, no caso do sexo masculino existem mais dados do que no sexo feminino e que as faixas etárias que apresentam uma maior quantidade de dados, em ambos os sexos, são referentes aos 40-65 anos, que por sua vez é a faixa etária onde se observam mais semelhanças tanto na Figura 8.107 como na Figura 8.109.

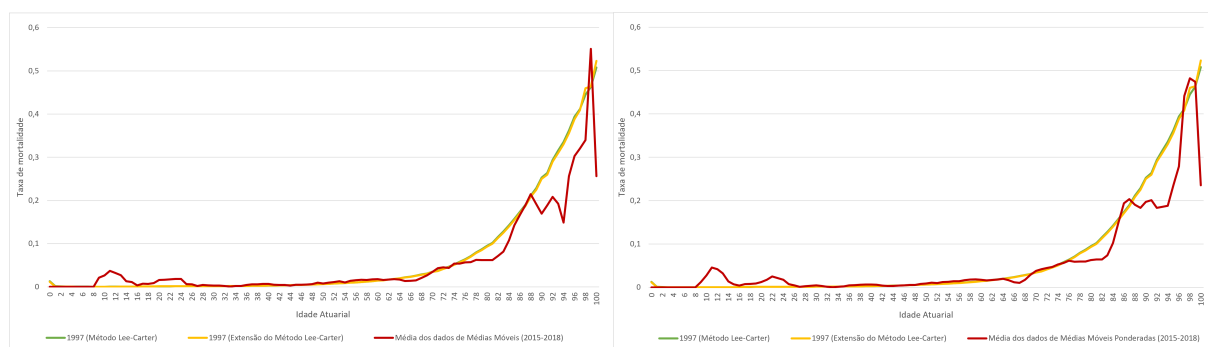


Figura 8.106: Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa - Sexo Masculino.



Figura 8.107: Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa entre os 25 e 65 anos - Sexo Masculino.

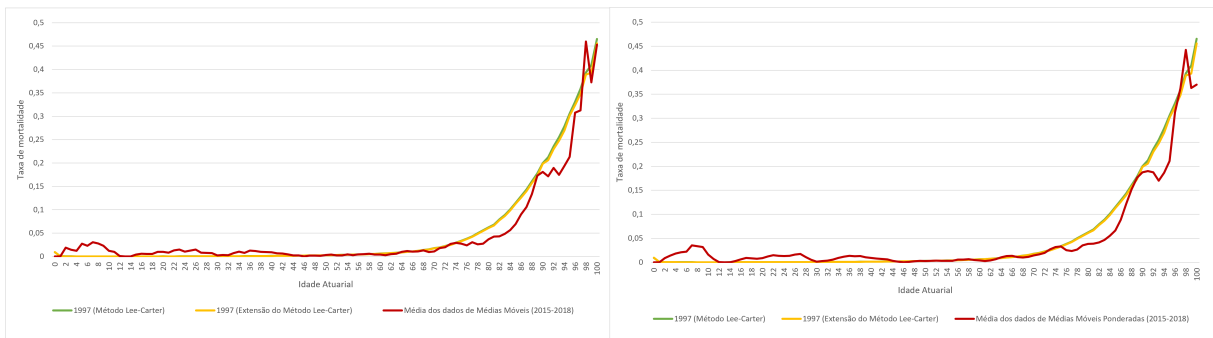


Figura 8.108: Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa - Sexo Feminino.

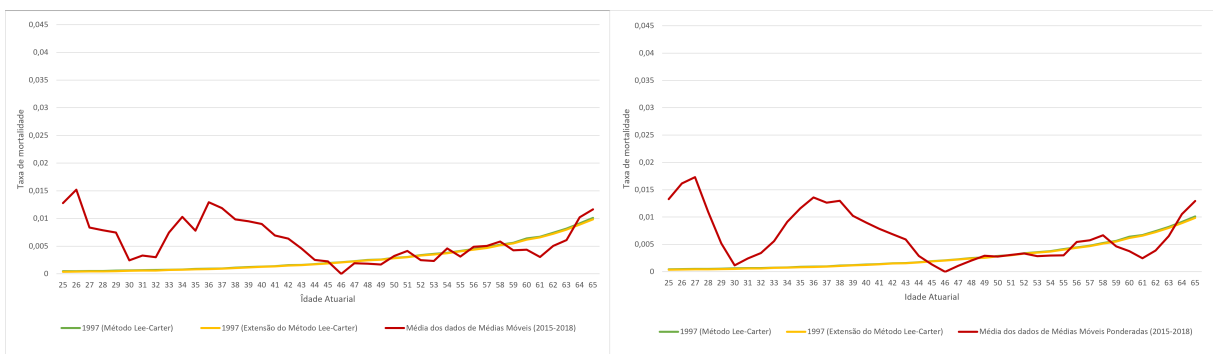


Figura 8.109: Comparação entre a média das taxas de mortalidade de 2015-2018 da população da companhia e as taxas de mortalidade de 1997 da população portuguesa entre os 25 e 65 anos - Sexo Feminino.

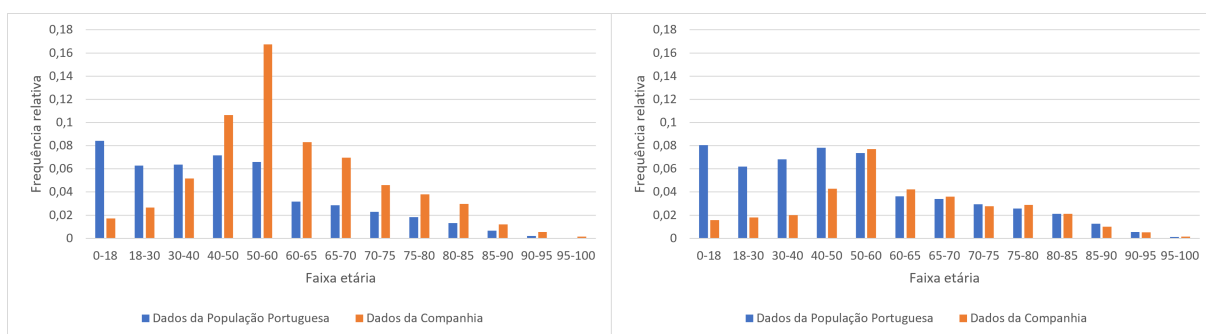


Figura 8.110: Comparação da frequência relativa dos dados da exposição ao risco da população portuguesa e da população da companhia, no sexo masculino (à esquerda) e no sexo feminino (à direita).

Para realizar uma melhor análise da comparação das curvas, nas Tabelas G.1, G.2, G.3, G.4, G.5, G.6, G.7 e G.8 do Apêndice G estão apresentadas as medidas de avaliação calculadas por faixas etárias, utilizando tanto as taxas de mortalidade da população da companhia calculadas através de Médias Móveis como de Médias Móveis Ponderadas e as taxas de mortalidade da população portuguesa obtidas pelo Método Lee-Carter e pela Extensão do Método Lee-Carter, com distinção de sexo. Com base nas mesmas constata-se que, tanto no sexo masculino como no sexo feminino, as medidas de avaliação são superiores a partir dos 75 anos. Relativamente à faixa etária que apresenta menores medidas de avaliação, observa-se que no sexo masculino corresponde à faixa dos 40-50 anos, já no sexo feminino corresponde aos 50-60 anos. Através destas tabelas observa-se também que as medidas de avaliação apresentam valores mais baixos quando calculadas entre as taxas de mortalidade da companhia e as taxas da população portuguesa obtidas pela Extensão do Método Lee-Carter.

No que se refere ao tipo de metodologia utilizada nos dados da companhia, verifica-se que, no caso do sexo masculino, as medidas calculadas com as taxas de mortalidade obtidas através das Médias Móveis apresentam valores inferiores quando comparadas com as calculadas a partir das taxas obtidas pelas Médias Móveis Ponderadas. Porém, no sexo feminino, constata-se o contrário, ou seja, as medidas de avaliação são mais reduzidas quando a comparação é realizada com os dados da companhia obtidos pelas Médias Móveis Ponderadas. De realçar que as medidas obtidas são muito semelhantes comparando ambas as metodologias. Por tudo o que foi referido, pode-se concluir que o comportamento dos pensionistas é semelhante ao comportamento da população portuguesa.

9 Conclusão

Este trabalho centrou-se na construção de tábuas de mortalidade da população portuguesa e no cálculo de taxas vitalícias dos pensionistas, mais precisamente no ramo de acidentes de trabalho. A construção das tábuas de mortalidade foi realizada através de dados entre os anos 1940-2019 referentes ao número de mortes e ao número de expostos ao risco, com e sem distinção de sexo. Para o cálculo de todas as funções que constituem uma tábua de mortalidade é necessário, inicialmente, ajustar a taxa de mortalidade de cada idade e para isso foram utilizados dois métodos, o Método Lee-Carter e o Modelo de Poisson (ou Extensão do Método Lee-Carter).

Apesar do objetivo deste trabalho não ser a previsão de dados, como trabalho complementar foram utilizados dois tipos de modelação para modelar e prever o nível geral de mortalidade obtido através do Método Lee-Carter e do Modelo de Poisson, de modo a obter um modelo que explicasse melhor os dados e que pudesse ser utilizado pela empresa para realizar previsões para anos futuros. Para isso foi utilizada a modelação ARIMA e o Método de Holt. Depois de comparados os modelos obtidos através das duas metodologias, constata-se que apesar das taxas de cobertura dos intervalos de previsão serem iguais a 100% em todos os casos, os modelos produzidos pela modelação ARIMA são mais precisos do que os produzidos pelo Método de Holt. Por essa razão será mais vantajoso utilizar a modelação ARIMA para realizar previsões do nível geral de mortalidade e, conseqüentemente, da taxa de mortalidade.

Depois de aplicados os dois métodos e calculadas as restantes funções da tábua de mortalidade, constata-se que a esperança de vida calculada através da taxa de mortalidade ajustada pelo Método Lee-Carter é menor na maior parte das idades e dos anos, comparativamente com a calculada através da Extensão do Método Lee-Carter. Neste caso o método preferível é aquele que produz uma esperança de vida menor, uma vez que isso irá resultar em menores taxas vitalícias, ou seja, numa menor pensão que será paga ao pensionista no caso de este ter sofrido um acidente de trabalho. Contudo, é necessário realçar que as diferenças existentes entre as esperanças de vida obtidas pelos dois métodos não são muito expressivas.

Como foi referido ao longo do trabalho, um dos objetivos era efetuar um ajustamento de tudo aquilo que foi referido anteriormente aos dados específicos do mercado segurador. Para isso, com os dados fornecidos pela empresa, foram calculadas as taxas de mortalidade de cada idade entre os anos 2007 e 2021. Visto que os dados fornecidos apresentavam bastantes idades com valores nulos, decidiu-se utilizar as metodologias de Médias Móveis e Médias Móveis Ponderadas para contornar esta situação. Depois de transformados os dados originais, foram calculadas as taxas de mortalidade e, de forma a entender se o comportamento da população da companhia era semelhante ao da população portuguesa, foram calculadas cinco medidas de avaliação e através destas observam-se conclusões diferentes relativamente ao sexo masculino e feminino. No caso do sexo masculino, verificam-se valores menores das respetivas medidas quando estas são calculadas com as taxas de mortalidade da companhia obtidas pelas Médias Móveis e as taxas da população portuguesa obtidas pela Extensão do Método Lee-Carter. Já no sexo feminino as medidas de avaliação são menores quando são comparadas as taxas de mortalidade da companhia obtidas pelas Médias Móveis Ponderadas e as taxas da população portuguesa obtidas pela Extensão do Método Lee-Carter.

Em suma, relativamente à utilização do Método Lee-Carter ou à sua Extensão para o ajustamento da taxa de mortalidade, devido ao Método Lee-Carter não conseguir prever acontecimentos bruscos que possam acontecer no futuro, como por exemplo, o aparecimento de algum vírus ou doença, alguma catástrofe natural, faz com que a Extensão do Método Lee-Carter seja o método mais indicado. Outra vantagem da Extensão do Método LC face ao Método LC, que o torna mais atrativo em aplicações atuariais, é o facto de este ser um método flexível, isto é, este pode ser expandido de modo a considerar o impacto sobre a mortalidade de outras variáveis explicativas que não apenas a idade ou o tempo cronológico, por exemplo, a profissão, o grau de incapacidade, o nível de rendimento, entre outros.

9.1 Sugestões para Trabalho Futuro

Nesta secção são evidenciadas algumas questões que surgiram depois de realizar este trabalho e que podem ser estudadas numa futura investigação.

Ao realizar a análise dos dados, tanto da população portuguesa como da companhia, foi notória a presença de vários *outliers*, mais especificamente nos dados referentes ao número de mortes. Apesar de estes não serem considerados erros mas sim valores extremos, seria indicado realizar o tratamento dos mesmos, de modo a reduzir o ruído que estes provocam nos dados. Outra questão importante recai sobre

a longevidade dos sinistrados, mais concretamente à presença ou não de incapacidade. Deste modo, seria adequado estudar a longevidade dos sinistrados consoante a incapacidade que estes apresentassem, de forma a comparar o comportamento da população "saudável" (i.e. sem qualquer tipo de incapacidade) e dos incapacitados.

Relativamente aos dados da companhia e às metodologias adotadas para combater os vários valores nulos presentes nos mesmos, propõe-se a utilização de outros métodos de agregação de dados, de modo a entender se os métodos utilizados usualmente pela empresa, ou seja, o método de Médias Móveis e Médias Móveis Ponderadas, são os mais adequados ou não para contornar este problema.

Como foi referido ao longo do trabalho, a esperança de vida tem vindo a aumentar ao longo dos anos e esta tem tendência a aumentar nos próximos anos. Por esta razão, os resultados obtidos neste trabalho são viáveis a curto prazo, sendo por isso necessário realizar este estudo anualmente, de forma a que as tábuas de mortalidade estejam sempre atualizadas. Relativamente à utilização do Método Lee-Carter e do Modelo de Poisson, para o ajustamento da taxa de mortalidade, é necessário averiguar se a dimensão dos dados utilizada neste estudo para a aplicação dos dois métodos é a mais indicada ou não, ou seja, se é necessário truncar os dados em alguns anos de acordo com os dados existentes da companhia ou se podem ser utilizados todos os dados disponíveis como foi realizado neste estudo.

Por fim, em relação à comparação das curvas da taxa de mortalidade da população portuguesa e da população da companhia, no presente trabalho foram calculadas algumas medidas de avaliação de modo a entender se o comportamento de ambas as curvas era semelhante. Outra forma de realizar esta comparação seria através da aplicação de modelos relacionais, sendo que estes procuram estabelecer uma relação entre o comportamento de duas variáveis, que neste caso seriam a mortalidade da população portuguesa e da população da companhia. Outra questão que também poderia ser estudada era a análise e o tratamento quer da cauda inferior quer da superior dos dados, uma vez que nas caudas se observam mais dissemelhanças quando são comparadas as duas curvas da taxa de mortalidade referidas anteriormente.

Bibliografia

- Akaike, H. (1974). A new look at the statistical model identification. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected papers of hirotugu akaike* (p. 215–222). New York, NY: Springer New York.
- Alpuim, T. (1998). *Séries temporais* (2ª ed.). Associação dos Estudantes da Faculdade de Ciências de Lisboa.
- Alpuim, T., & Ribeiro, I. (2003, Apr). A state space model for rub-off triangles. *Applied Stochastic Models in Business and Industry*, 19, 105–120.
- Armstrong, J. S. (2001). *Principles of forecasting: A handbook for researchers and practioners*. (1ª ed.). Norwell, MA: Klumer.
- ASF. (2015). *Guia de seguros e fundos de pensões* (3ª ed.). Autoridade de Supervisão de Seguros e Fundos de Pensões.
- Barreto, X., Correia, J. P., Cunha, O., Matos, A., Peixoto, J., Machado, J. C., ... Santos, N. S. (2014). *A mortalidade infantil em portugal: evolução dos indicadores e factores associados entre 1988 a 2008*. Fundação Francisco Manuel dos Santos.
- Batista, R. (2021). *Mercado segurador* (Unpublished doctoral dissertation). Instituto Superior de Contabilidade e Administração de Coimbra.
- Bell, W. (1997). Comparing and assessing time series methods for forecasting age specific demographic rates. *Journal of Official Statistics*(13), 279–303.
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: forecasting and control* (5ª ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.

- Bravo, J. (2007). *Tábuas de mortalidade contemporâneas e prospectivas: Modelos estocásticos, aplicações actuariais e cobertura do risco de longevidade* (Unpublished doctoral dissertation). Universidade de Évora.
- Brouhns, N., Denuit, M., & Vermunt, J. (2002, 12). A poisson log-bilinear regression approach to the construction of projected life tables. *Insurance: Mathematics and Economics*, 31, 373-393.
- Brown, R. (1963). *Smoothing, forecasting and prediction of discrete time series*. NJ:Prentice-Hall.
- Caiado, J. (2022). *Métodos de previsão em gestão - com aplicações em excel* (3ª ed.). Edições Sílabo.
- Carrilho, M. J., & Patrício, L. (2004). *Revista de estudos demográficos n°36*. Instituto Nacional de Estatística.
- Carter, L., & Prskawetz, A. (2001). *Examining structural shifts in mortality using the lee-carter method*. (Unpublished doctoral dissertation). Max Planck Institute for Demographic Research WP 2001- 007, Germany.
- Carter, L. R. (1996, Jan). Forecasting u.s. mortality: A comparison of box-jenkins arima and structural time series models. *The Sociological Quarterly*, 37(1), 127–144.
- Chatfield, C. (2000). *Time-series forecasting*. Chapman and Hall/CRC.
- Coelho, E. (2005). *Revista de estudos demográficos n°37*. Instituto Nacional de Estatística.
- Coelho, E., & Malta, J. (2014, Jan). *Tábuas completas de mortalidade*.
- Cordeiro, C. M. H. (2011). *Métodos de reamostragem em modelos de previsão*. (Unpublished doctoral dissertation). Instituto Superior de Agronomia.
- Cowpertwait, P. S. P., & Metcalfe, A. V. (2009). *Introductory time series with r*. Dordrecht New York: Springer.
- Database, H. M. (2022). *Reliability and accuracy matter*. <https://www.mortality.org/>. (Accessed: 2022-01-27)
- Dickey, D. A., & Fuller, W. A. (1979, Jun). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427–431.

- Enders, W. (2015). *Applied econometric time series* (4^a ed.). Hoboken, NJ: Wiley.
- Ferreira, A. I. (2012). *Projeções da taxa de mortalidade* (Unpublished doctoral dissertation). Faculdade de Ciências da Universidade do Porto.
- Gardner, E. S., & McKenzie, E. (1985, Oct). Forecasting trends in time series. *Management Science*, 31, 1237–1246.
- Holt, C. (1957). Forecasting trends and seasonal by exponentially weighted averages. *Office of Naval Research Memorandum*, 20.
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3^a ed.). Australia: OTexts.
- Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002, Jul). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, 18, 439–454.
- INE. (2020, Mar). *Projeções de população residente 2018-2080*. Instituto Nacional de Estatística.
- INE. (2021, May). *Tábuas de mortalidade para Portugal 2018-2020*. Instituto Nacional de Estatística.
- Lee, R. (2000, Jan). The lee-carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4(1), 80–91.
- Lee, R., & Miller, T. (2001). Evaluating the performance of the lee-carter method for forecasting mortality. *Demography*, 38(4), 537–549.
- Lee, R. D., & Carter, L. R. (1992, Sep). Modeling and forecasting u. s. mortality. *Journal of the American Statistical Association*, 87(419), 659.
- Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: methods and applications* (3^a ed.). New York: John Wiley & Sons.
- Menezes, R. (2020). *Sebenta de séries temporais*.
- Ng, S., & Perron, P. (1995, Mar). Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90, 268–281.

- Russolillo, M. (2014). The future human lifespan: A study on italian population. *Applied Mathematics*, 1641–1650.
- Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71, 599–607.
- Schwarz, G. (1978, Mar). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461–464.
- Schwert, G. W. (2002, Jan). Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics*, 20, 5–17.
- Shumway, R. H., & Stoffer, D. S. (2011). *Time series analysis and its applications: with r examples* (3^a ed.). New York: Springer.
- Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19, 715-725.
- Wang, S. (2006). *Exponential smoothing for forecasting and bayesian validation of computer models* (Unpublished doctoral dissertation). Industrial and Systems Engineering Georgia Institute of Technology.
- Wilmoth. (1993). *Computational methods for fitting and extrapolating the leecarter model of mortality change*. (Techn'ical Report). Department of Demography, University of California, Berkeley.
- Wilmoth, J., & Valkonen, T. (2002). A parametric representation of mortality differentials over age and time. Pontignano, Italy: Fifth seminar of the EAPS Working Group on Differentials in Health, Morbidity and Mortality in Europe.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324-342.

Apêndice A

Análise Descritiva dos Dados Portugueses

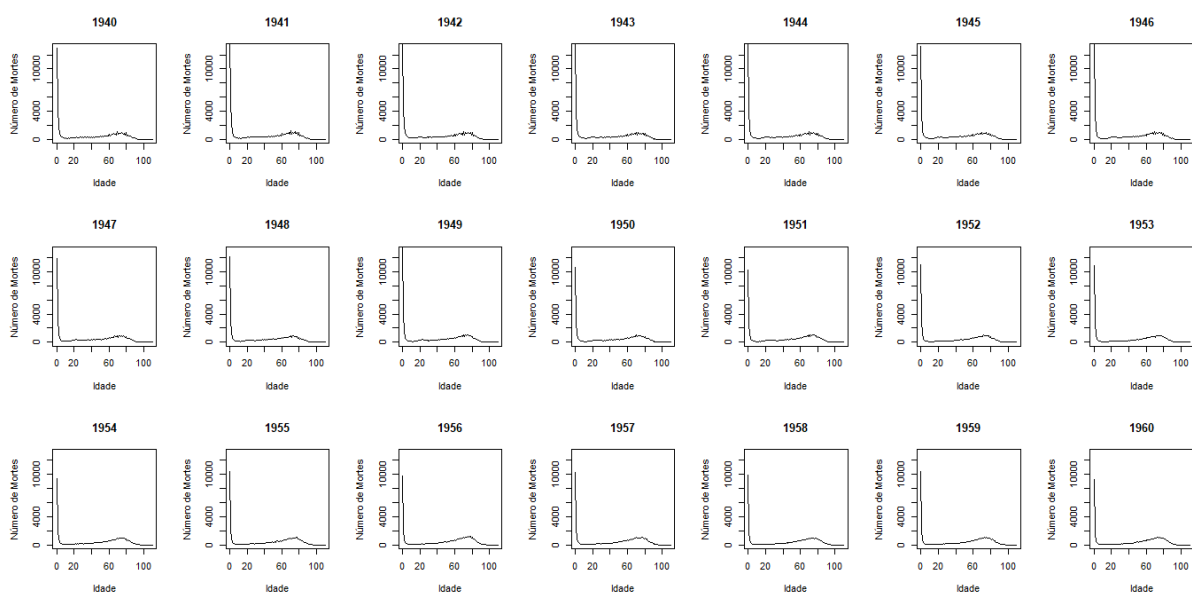


Figura A.1: Número de Mortes nos Anos 1940-1960, para o Sexo Masculino.

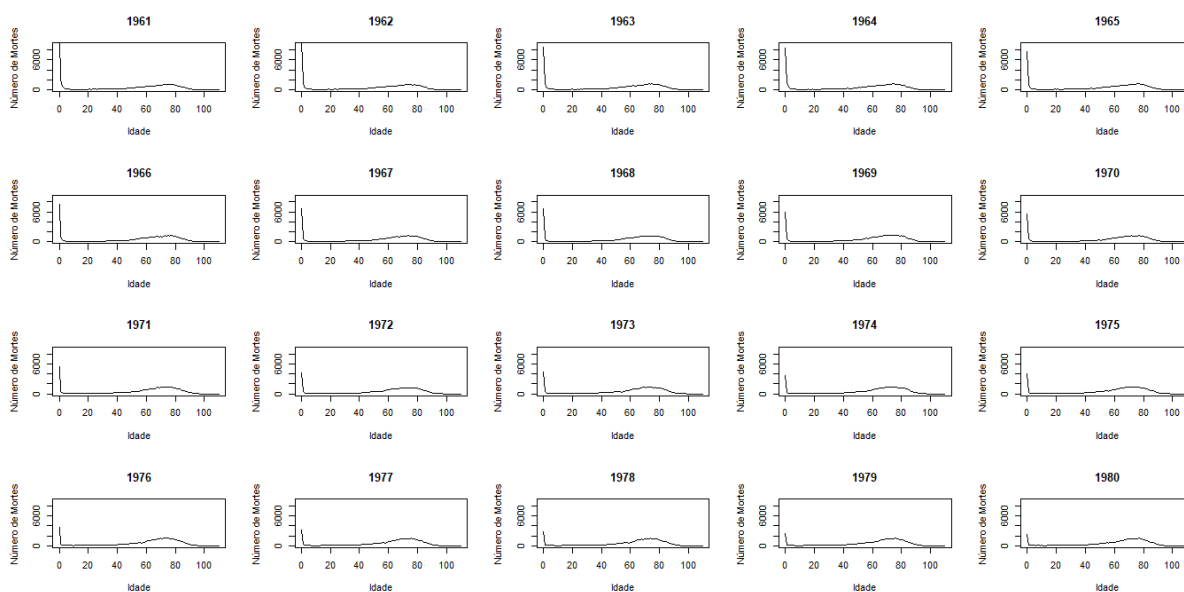


Figura A.2: Número de Mortes nos Anos 1961-1980, para o Sexo Masculino.

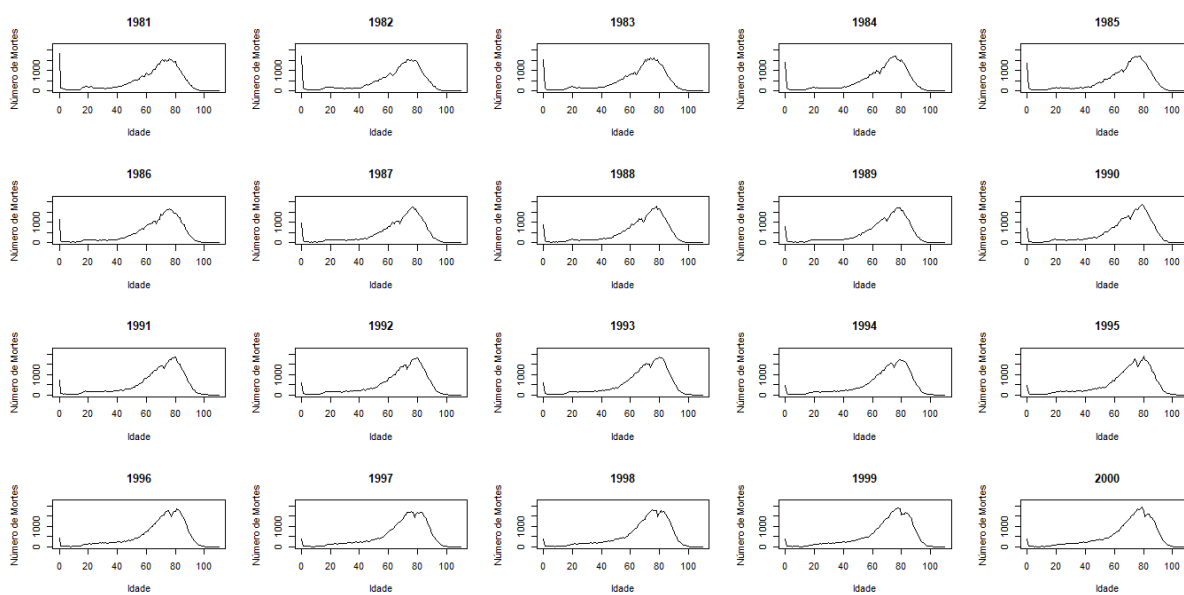


Figura A.3: Número de Mortes nos Anos 1981-2000, para o Sexo Masculino.

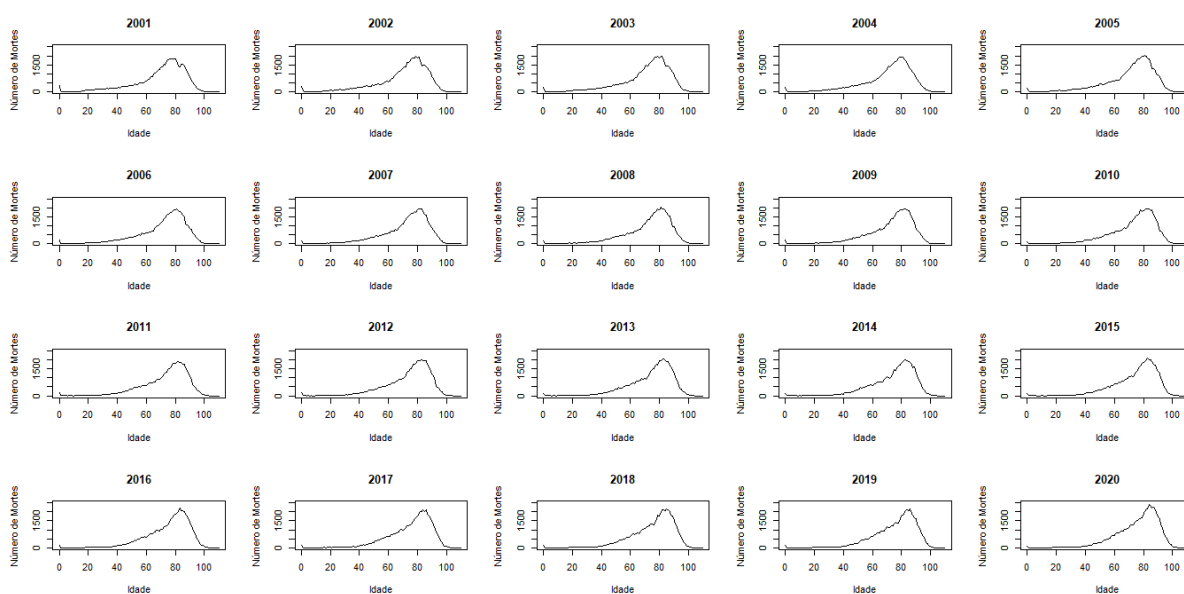


Figura A.4: Número de Mortes nos Anos 2001-2020, para o Sexo Masculino.

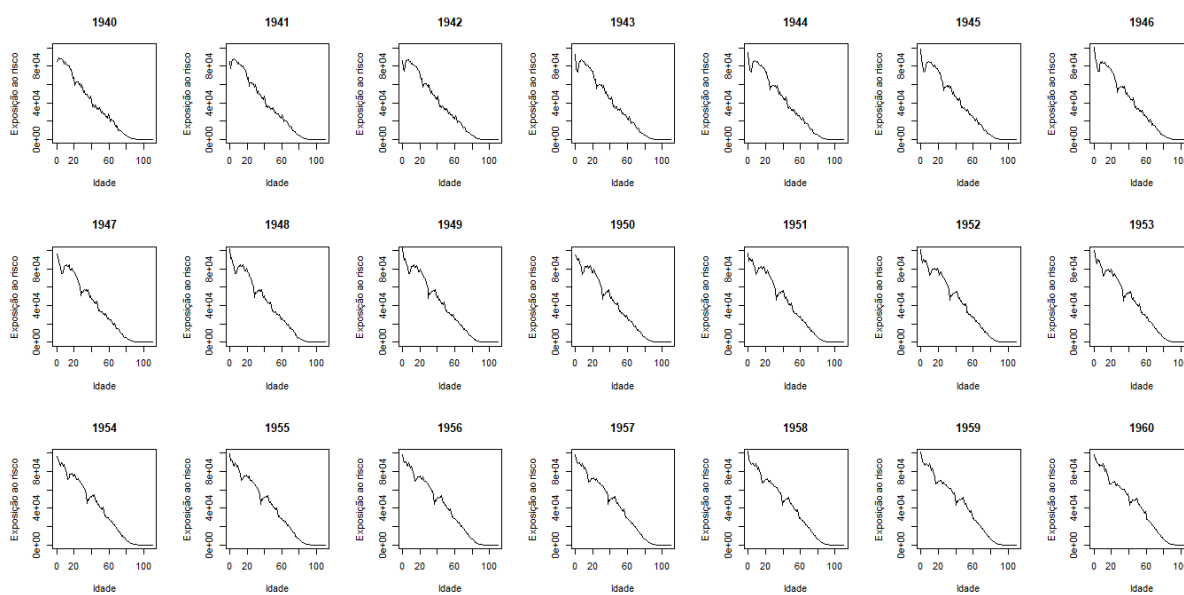


Figura A.5: Número de Expostos ao Risco nos Anos 1940-1960, para o Sexo Masculino.

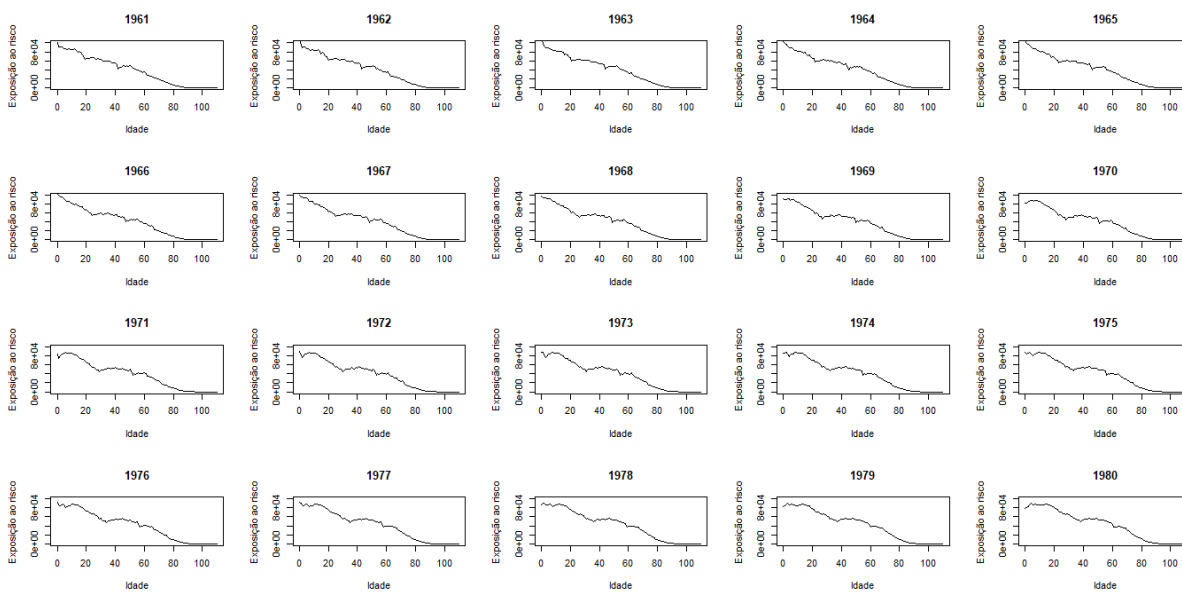


Figura A.6: Número de Expostos ao Risco nos Anos 1961-1980, para o Sexo Masculino.

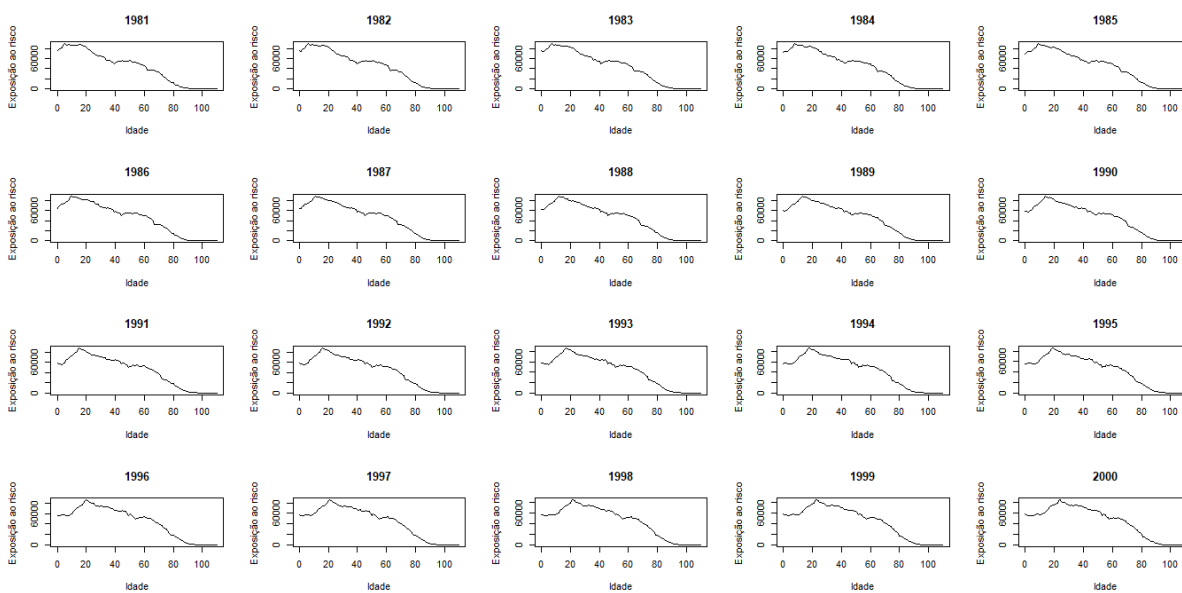


Figura A.7: Número de Expostos ao Risco nos Anos 1981-2000, para o Sexo Masculino.

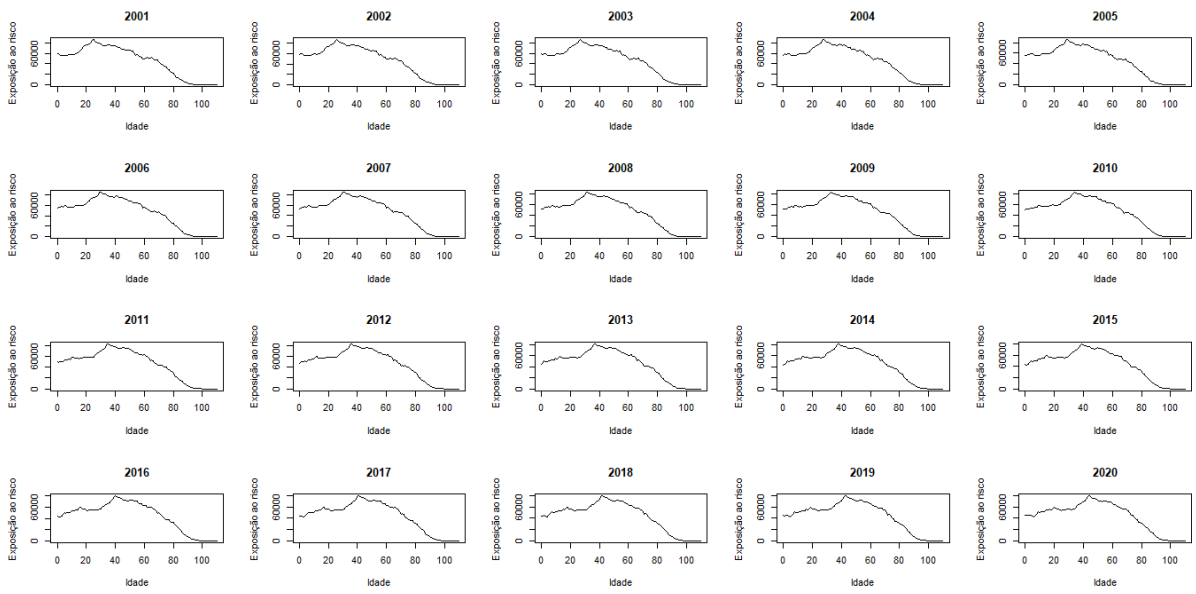


Figura A.8: Número de Expostos ao Risco nos Anos 2001-2020, para o Sexo Masculino.

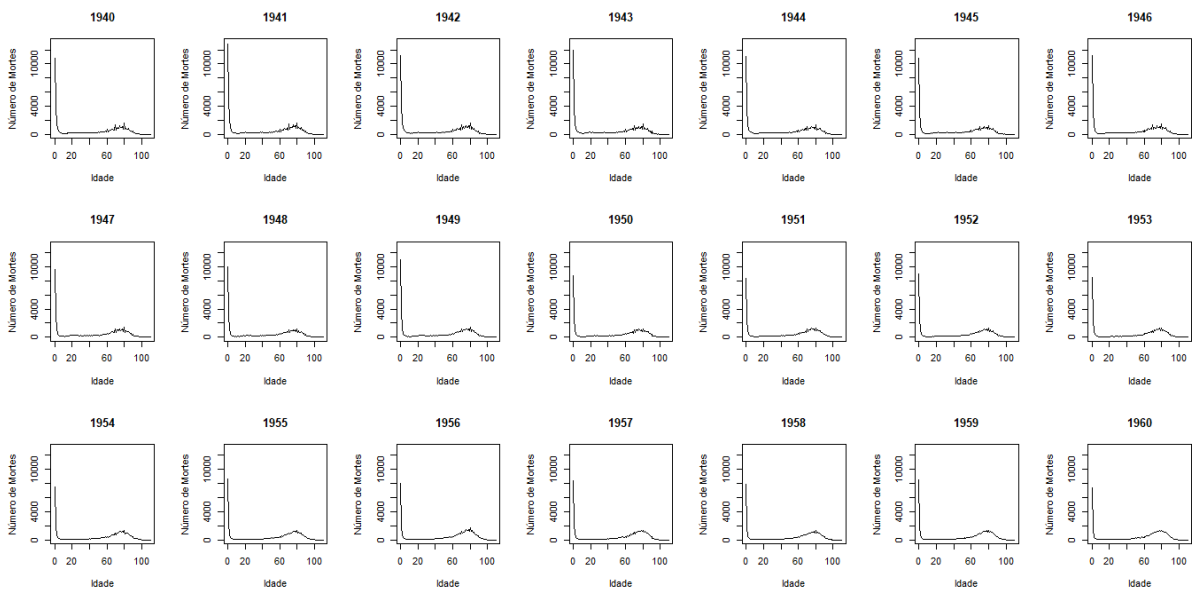


Figura A.9: Número de Mortes nos Anos 1940-1960, para o Sexo Feminino.

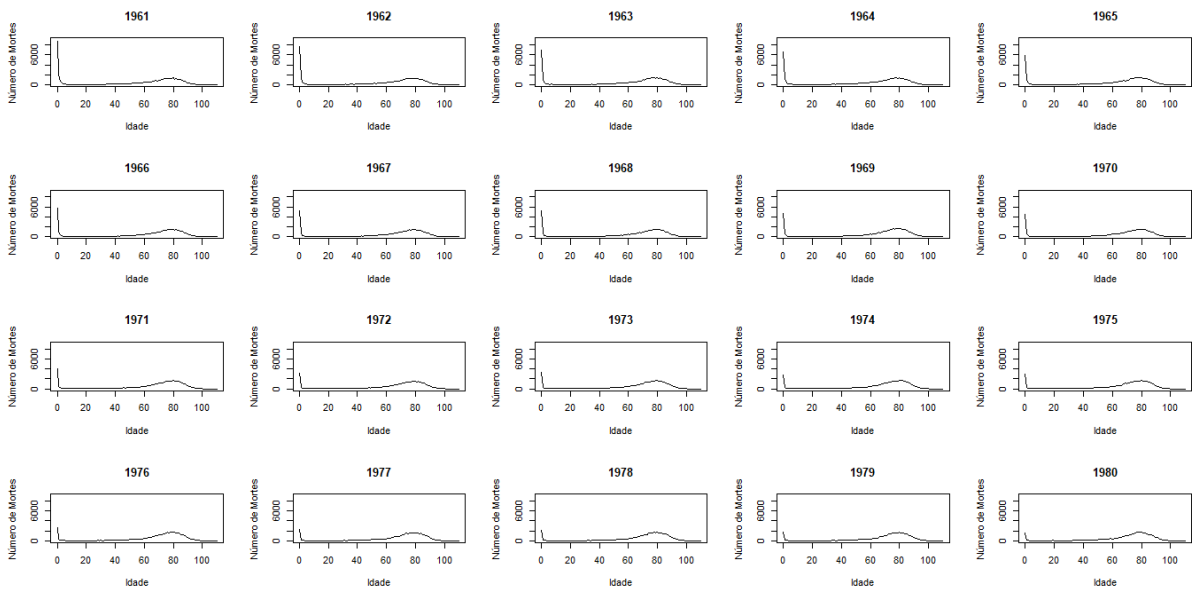


Figura A.10: Número de Mortes nos Anos 1961-1980, para o Sexo Feminino.

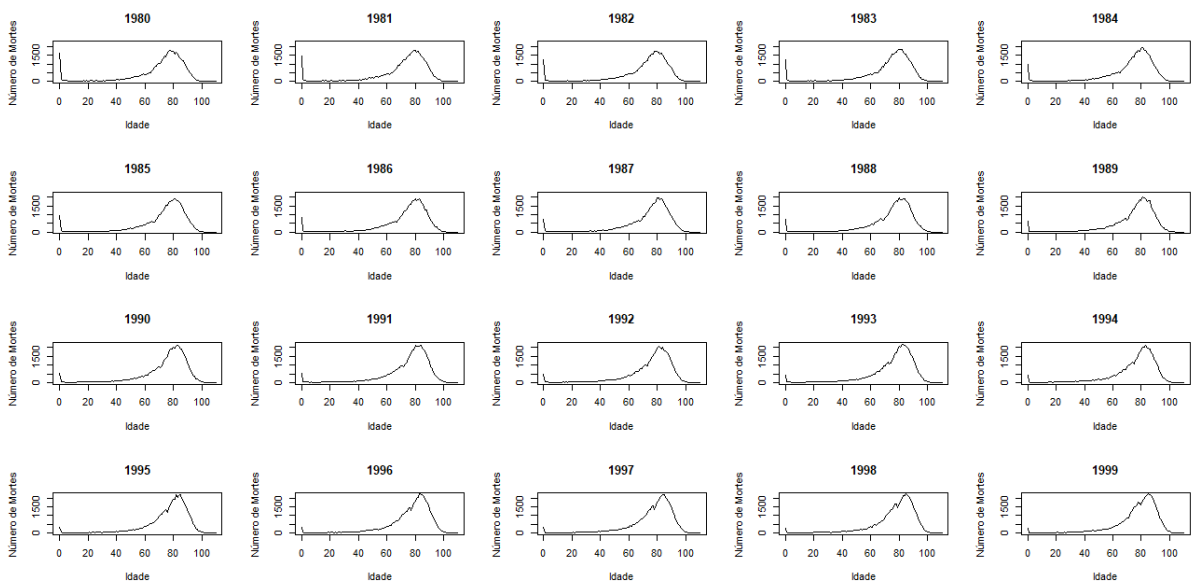


Figura A.11: Número de Mortes nos Anos 1981-2000, para o Sexo Feminino.

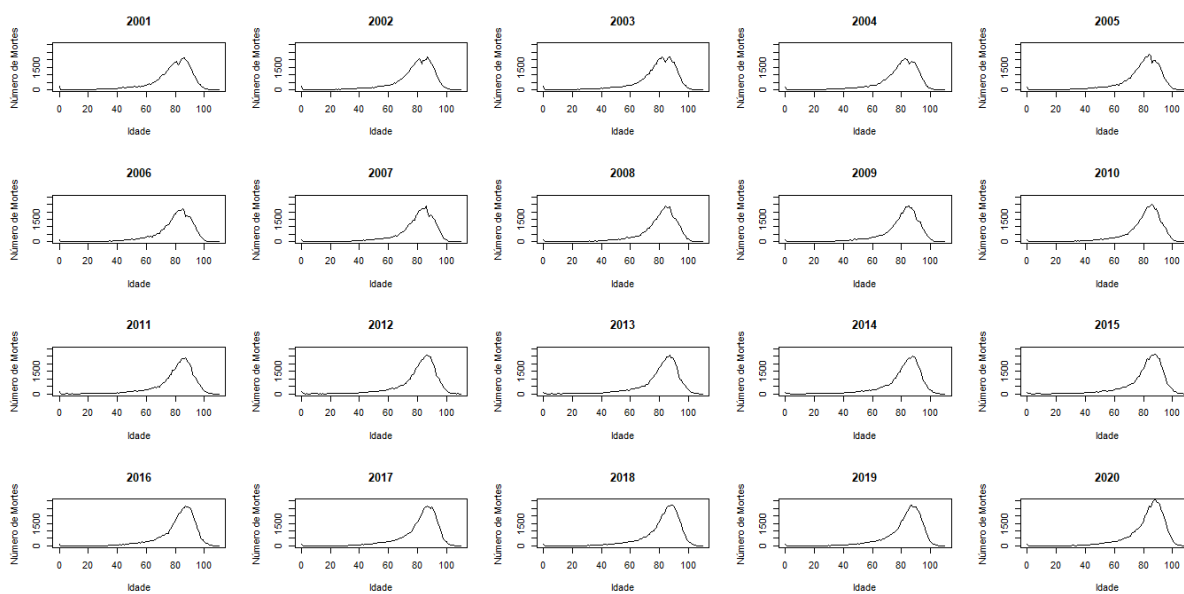


Figura A.12: Número de Mortes nos Anos 2001-2020, para o Sexo Feminino.

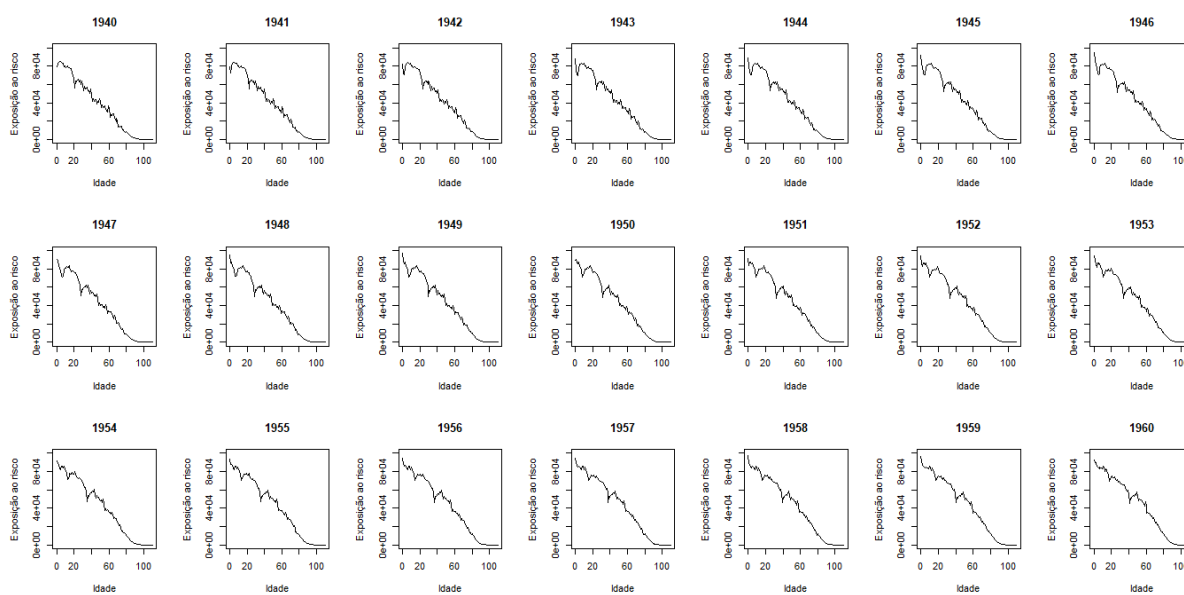


Figura A.13: Número de Expostos ao Risco nos Anos 1940-1960, para o Sexo Feminino.

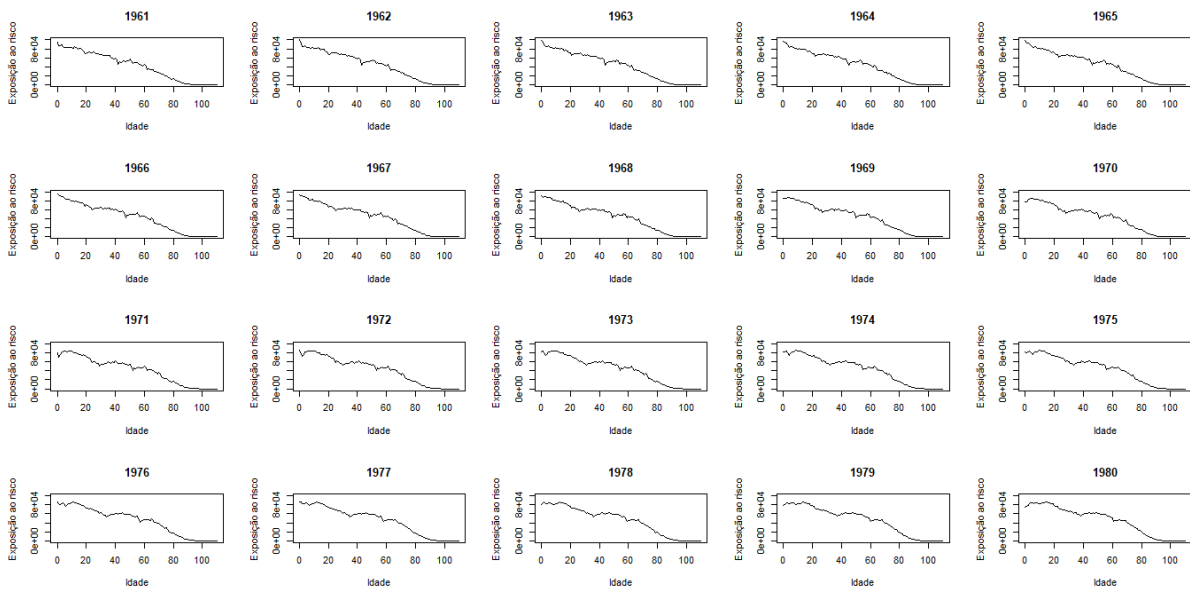


Figura A.14: Número de Expostos ao Risco nos Anos 1961-1980, para o Sexo Feminino.

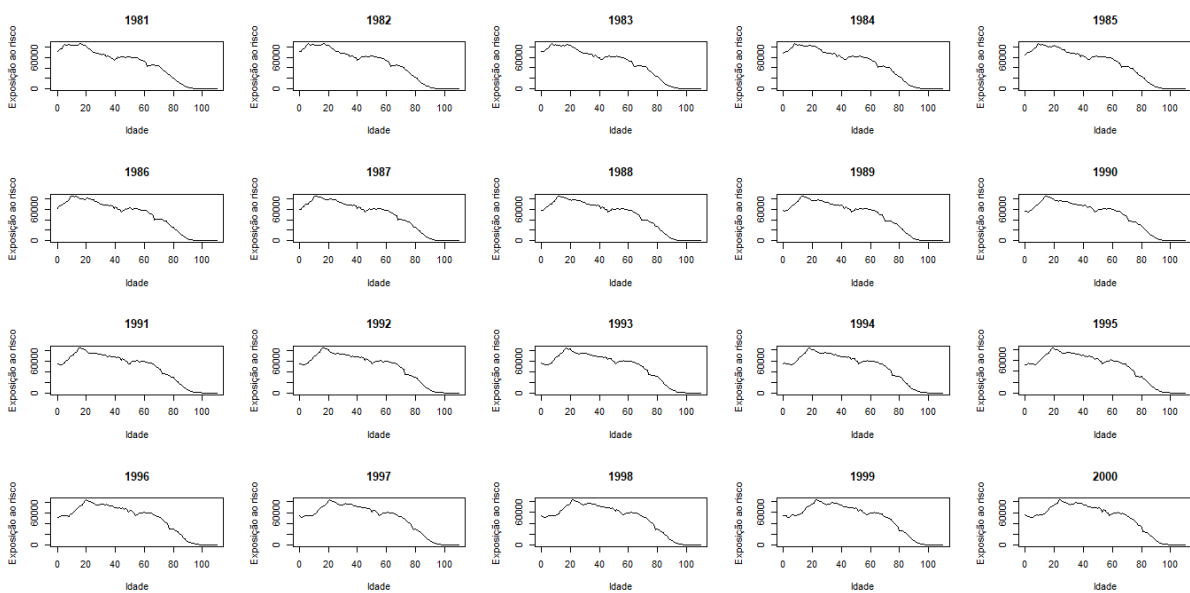


Figura A.15: Número de Expostos ao Risco nos Anos 1981-2000, para o Sexo Feminino.

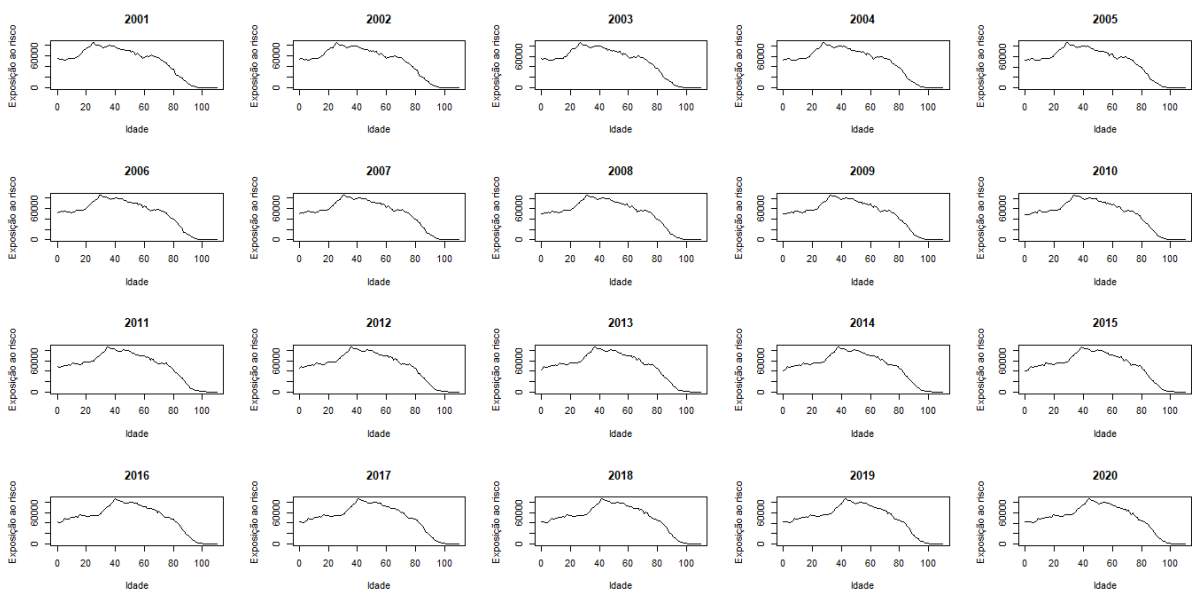


Figura A.16: Número de Expostos ao Risco nos Anos 2001-2020, para o Sexo Feminino.

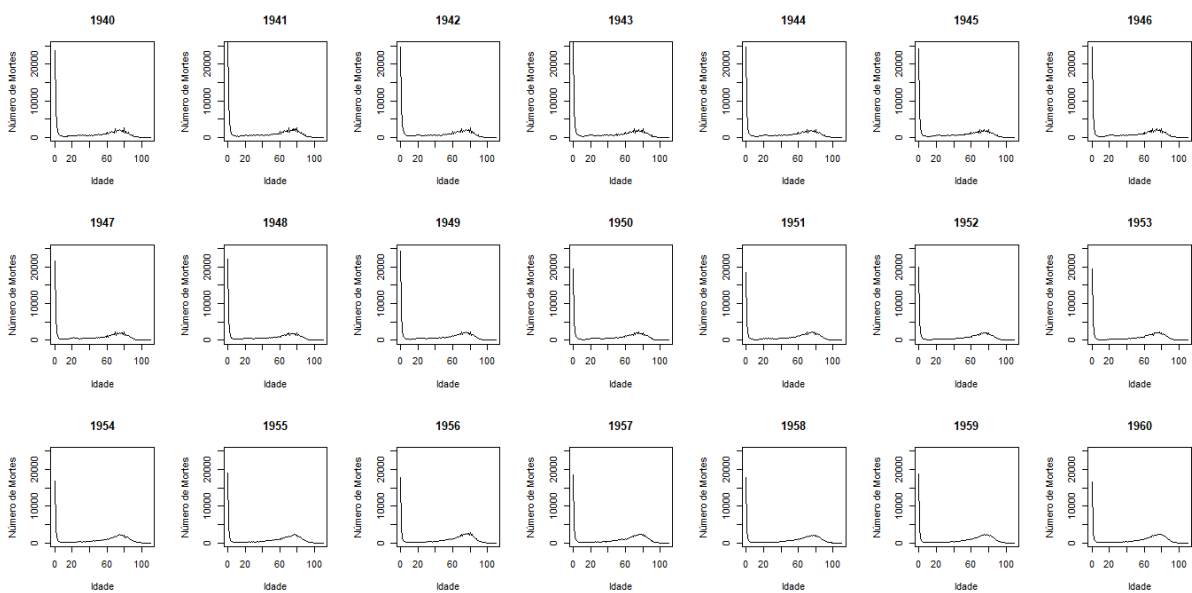


Figura A.17: Número de Mortes nos Anos 1940-1960, Sem Distinção de Sexo.

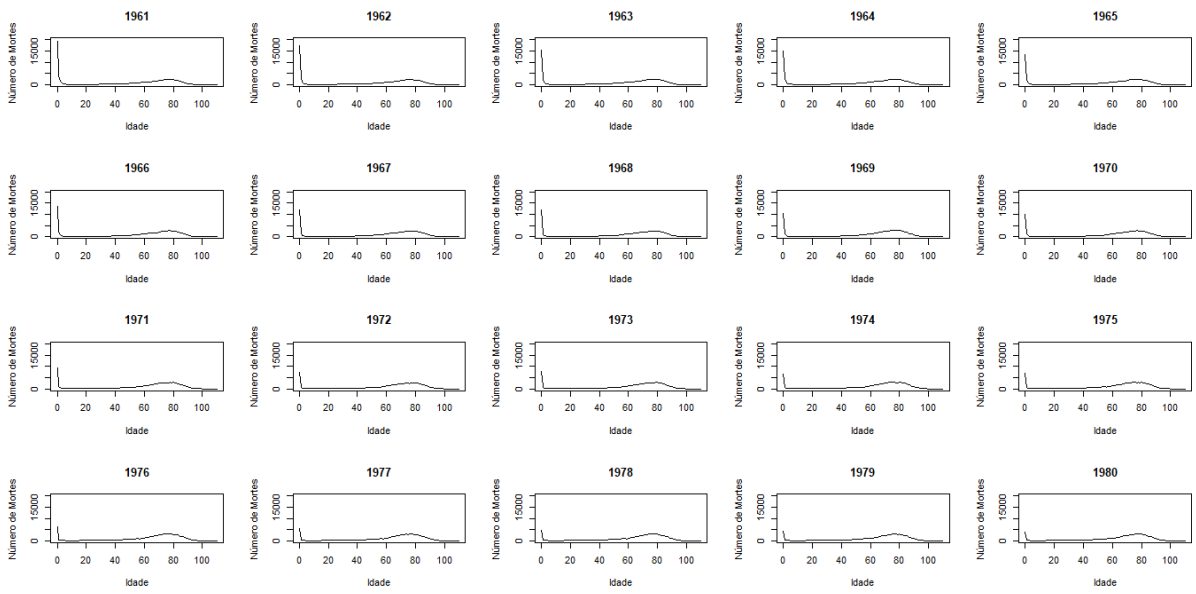


Figura A.18: Número de Mortes nos Anos 1961-1980, Sem Distinção de Sexo.

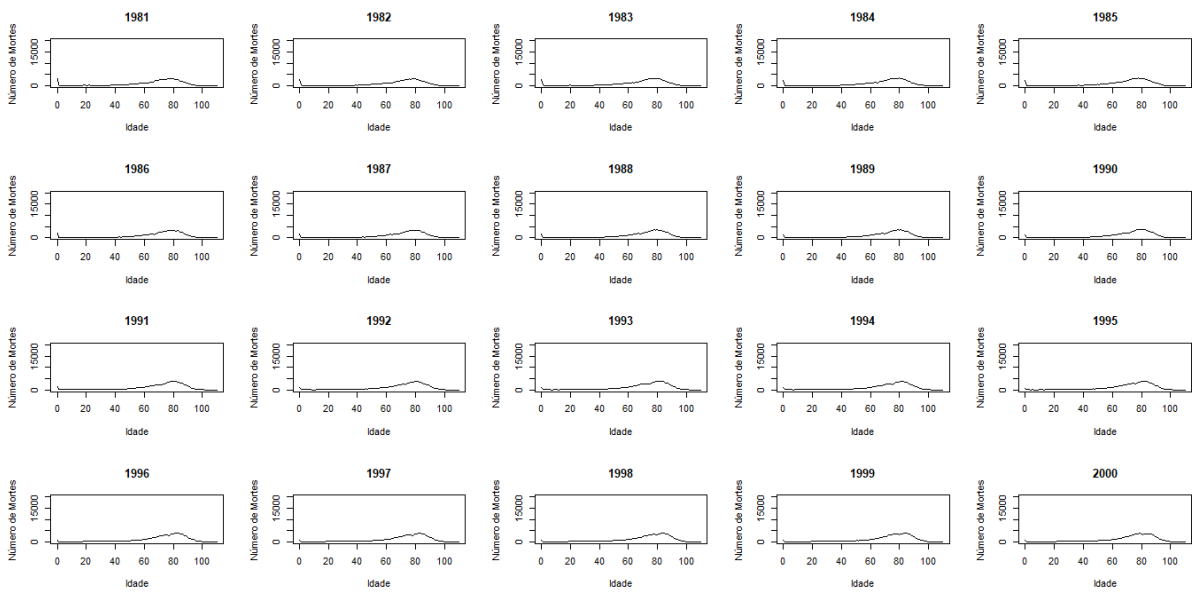


Figura A.19: Número de Mortes nos Anos 1981-2000, Sem Distinção de Sexo.

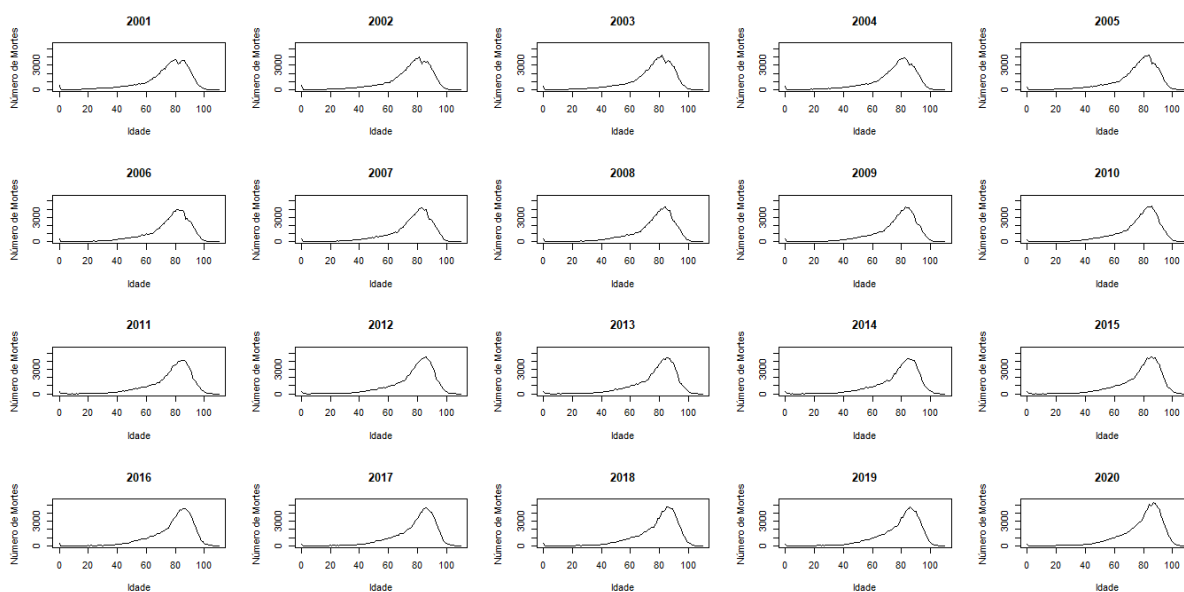


Figura A.20: Número de Mortes nos Anos 2001-2020, Sem Distinção de Sexo.

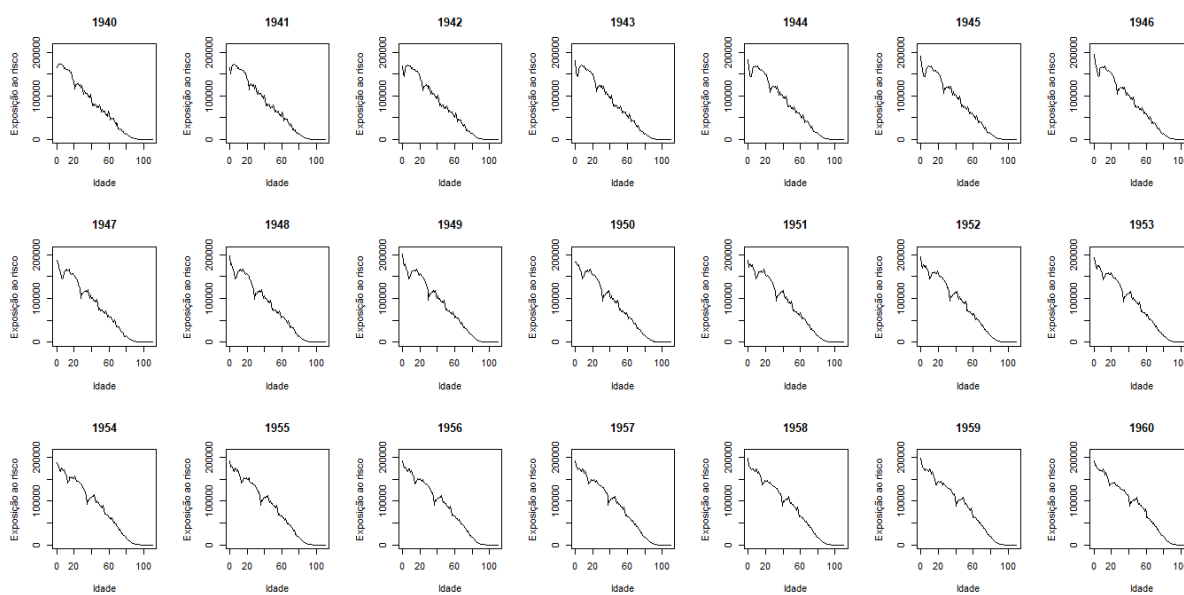


Figura A.21: Número de Expostos ao Risco nos Anos 1940-1960, Sem Distinção de Sexo.

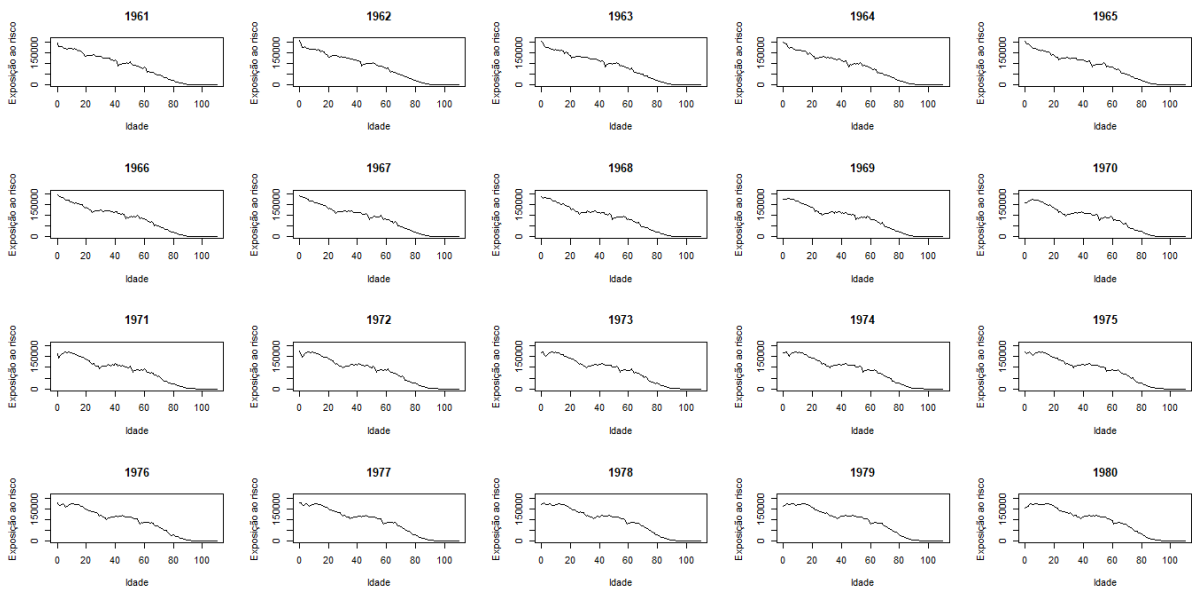


Figura A.22: Número de Expostos ao Risco nos Anos 1961-1980, Sem Distinção de Sexo.

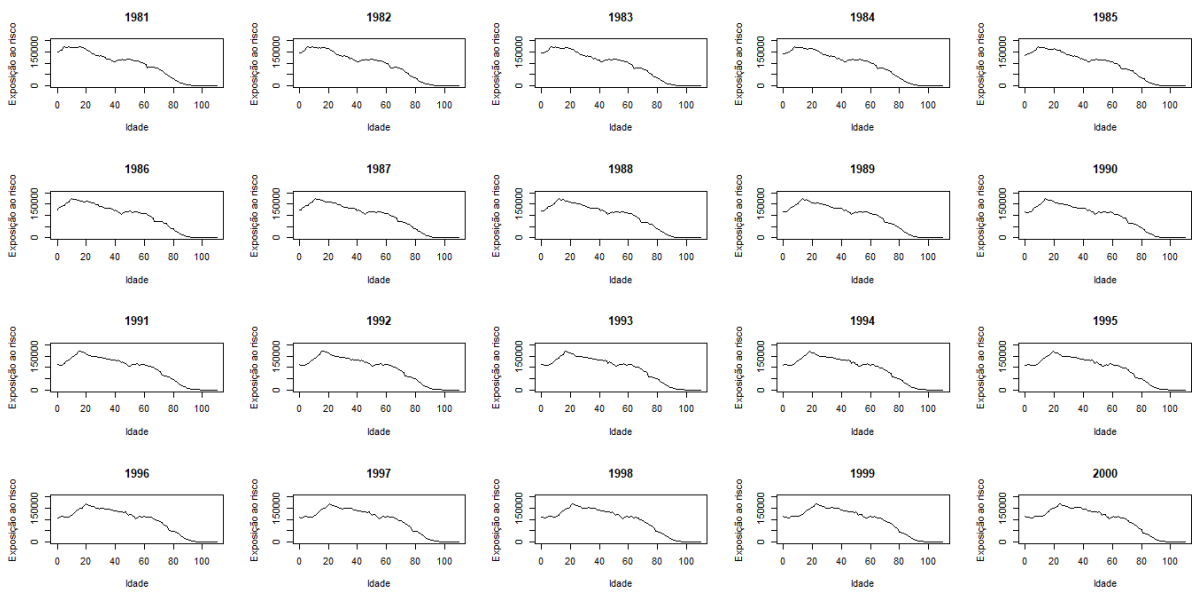


Figura A.23: Número de Expostos ao Risco nos Anos 1981-2000, Sem Distinção de Sexo.

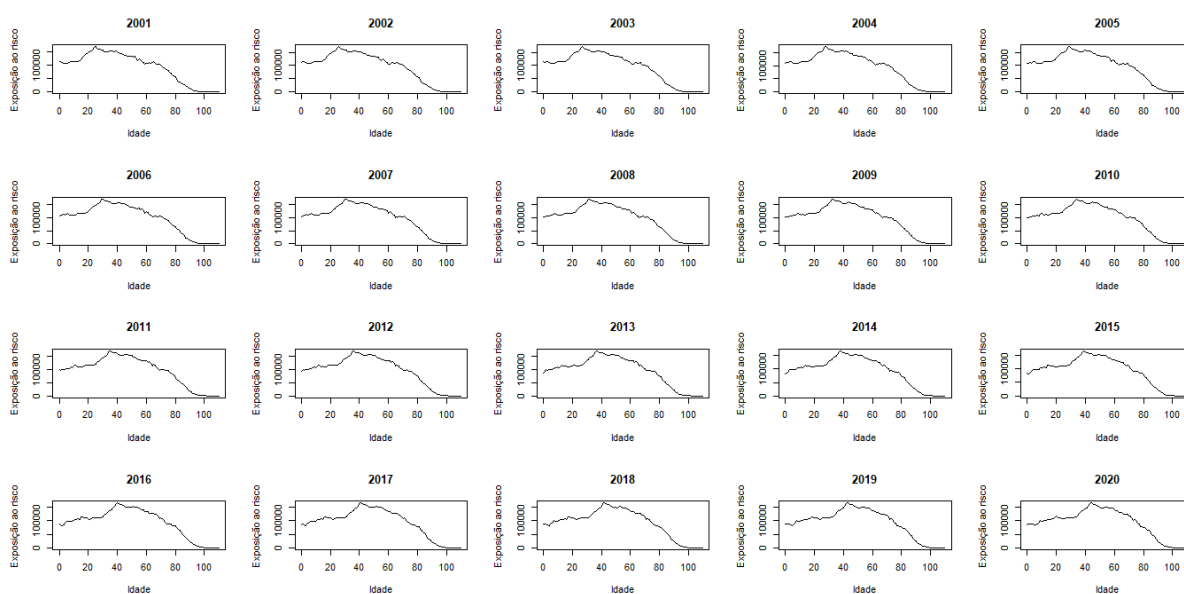


Figura A.24: Número de Expostos ao Risco nos Anos 2001-2020, Sem Distinção de Sexo.

Tabela A.1: Medidas descritivas dos dados de exposição ao risco dos três casos de estudo.

Medida Descritiva	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
início	1940	1940	1940
fim	2020	2020	2020
n° de observações	8991	8991	8991
mínimo	0	0	0
1° quartil	7475,265	14042,480	21800,660
média	40830,880	44371,570	85202,450
mediana	47921,460	52980,100	101711,530
3° quartil	64990,280	68740,450	133588,790
máximo	106015,900	100419,800	206435,670
desvio padrão	29681,170	29011,210	58520,760

Tabela A.2: Medidas descritivas dos dados de número de mortes dos três casos de estudo.

Medida Descritiva	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
início	1940	1940	1940
fim	2020	2020	2020
n° de observações	8991	8991	8991
mínimo	0	0	0
1° quartil	72,090	46,000	138,000
média	478,652	453,360	932,0126
mediana	233,220	177,000	430,000
3° quartil	681,890	597,275	1305,495
máximo	14965,900	12880,600	27846,500
desvio padrão	792,602	748,987	1516,955

Tabela A.3: Médias anuais (de 10 em 10 anos) dos dados de exposição ao risco dos três casos de estudo.

Ano	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
1940	33322,590	36095,250	69417,840
1950	36465,390	39386,420	75851,810
1960	38288,700	41743,180	80031,870
1970	37141,990	41087,410	78229,400
1980	42437,640	45736,350	88173,990
1990	43426,320	46728,390	90154,710
2000	44751,990	48042,500	92794,500
2010	45614,980	49763,670	95378,650
2020	43785,350	48997,120	92782,470

Tabela A.4: Médias anuais (de 10 em 10 anos) dos dados de número de mortes dos três casos de estudo.

Ano	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
1940	548,919	536,541	1085,459
1950	471,765	454,341	926,107
1960	434,325	421,595	855,919
1970	426,703	411,973	838,676
1980	446,225	407,775	854,000
1990	479,216	446,622	925,838
2000	495,703	453,523	949,225
2010	488,463	466,076	954,540
2020	552,882	558,450	1111,332

Apêndice B

Estimação dos Parâmetros (Método Lee-Carter)

Tabela B.1: Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sexo Masculino).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,749	0,028
1	-5,755	0,036
2	-6,355	0,032
3	-6,794	0,029
4	-7,048	0,027
5	-7,251	0,025
6	-7,457	0,026
7	-7,452	0,023
8	-7,610	0,023
9	-7,644	0,022
10	-7,667	0,020
11	-7,642	0,019
12	-7,667	0,019
13	-7,584	0,018
14	-7,398	0,017

Continuação da Tabela B.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
15	-7,185	0,015
16	-6,937	0,014
17	-6,804	0,015
18	-6,589	0,013
19	-6,499	0,012
20	-6,455	0,012
21	-6,424	0,012
22	-6,416	0,012
23	-6,401	0,012
24	-6,367	0,013
25	-6,351	0,012
26	-6,329	0,012
27	-6,308	0,012
28	-6,260	0,012
29	-6,260	0,011
30	-6,206	0,011
31	-6,194	0,011
32	-6,133	0,011
33	-6,077	0,011
34	-6,021	0,010
35	-5,960	0,010
36	-5,926	0,011
37	-5,866	0,010
38	-5,789	0,010
39	-5,733	0,009
40	-5,638	0,009
41	-5,612	0,008
42	-5,487	0,009
43	-5,437	0,008

Continuação da Tabela B.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
44	-5,358	0,008
45	-5,270	0,008
46	-5,222	0,007
47	-5,145	0,007
48	-5,036	0,007
49	-4,976	0,007
50	-4,872	0,007
51	-4,861	0,005
52	-4,735	0,006
53	-4,654	0,006
54	-4,592	0,006
55	-4,505	0,006
56	-4,434	0,007
57	-4,362	0,006
58	-4,269	0,006
59	-4,208	0,006
60	-4,075	0,007
61	-4,064	0,005
62	-3,932	0,007
63	-3,846	0,007
64	-3,769	0,007
65	-3,668	0,007
66	-3,590	0,007
67	-3,506	0,007
68	-3,392	0,007
69	-3,338	0,006
70	-3,171	0,008
71	-3,157	0,006
72	-2,996	0,008

Continuação da Tabela B.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
73	-2,909	0,007
74	-2,804	0,007
75	-2,683	0,008
76	-2,603	0,007
77	-2,492	0,007
78	-2,367	0,007
79	-2,317	0,006
80	-2,203	0,007
81	-2,185	0,005
82	-2,036	0,006
83	-1,949	0,005
84	-1,835	0,005
85	-1,732	0,005
86	-1,655	0,004
87	-1,574	0,004
88	-1,456	0,005
89	-1,415	0,003
90	-1,289	0,004
91	-1,317	0,001
92	-1,180	0,002
93	-1,127	0,001
94	-1,078	0,001
95	-0,992	0,001
96	-0,905	0,001
97	-0,879	0,000
98	-0,752	0,003
99	-0,762	0,000

100+	-0,706	-0,001
Fim da Tabela		

Tabela B.2: Estimação do Parâmetro k_t do Método Lee-Carter (Sexo Masculino).

Início da Tabela	
Ano	\hat{k}_t
1940	78,479
1941	84,694
1942	83,495
1943	79,933
1944	77,371
1945	74,183
1946	72,312
1947	66,047
1948	63,332
1949	65,133
1950	57,358
1951	55,605
1952	48,019
1953	40,606
1954	38,549
1955	38,216
1956	39,963
1957	39,220
1958	29,112
1959	32,953
1960	27,273
1961	31,421

Continuação da Tabela B.2

Ano	\hat{k}_t
1962	29,032
1963	28,077
1964	28,165
1965	23,267
1966	24,803
1967	23,383
1968	19,618
1969	21,583
1970	18,435
1971	22,591
1972	13,805
1973	16,264
1974	13,460
1975	16,847
1976	16,183
1977	9,794
1978	8,587
1979	5,164
1980	5,478
1981	3,505
1982	0,160
1983	-0,613
1984	-1,517
1985	-3,870
1986	-7,176
1987	-8,181
1988	-6,239
1989	-9,139
1990	-9,606

Continuação da Tabela B.2

Ano	\hat{k}_t
1991	-5,638
1992	-10,232
1993	-12,263
1994	-18,423
1995	-17,013
1996	-18,051
1997	-21,420
1998	-21,628
1999	-26,784
2000	-29,969
2001	-33,203
2002	-36,291
2003	-43,066
2004	-46,282
2005	-49,055
2006	-52,580
2007	-60,348
2008	-66,279
2009	-65,421
2010	-68,957
2011	-70,554
2012	-77,505
2013	-77,088
2014	-80,565
2015	-80,297
2016	-87,920
2017	-83,003
2018	-90,454
2019	-84,231

Continuação da Tabela B.2	
Ano	\hat{k}_t
2020	-30,608
Fim da Tabela	

Tabela B.3: Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sexo Feminino).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,937	0,023
1	-5,881	0,031
2	-6,573	0,028
3	-6,975	0,024
4	-7,282	0,023
5	-7,512	0,021
6	-7,589	0,019
7	-7,765	0,019
8	-7,873	0,018
9	-7,942	0,017
10	-7,950	0,015
11	-7,998	0,016
12	-8,010	0,016
13	-7,954	0,016
14	-7,835	0,016
15	-7,682	0,014
16	-7,609	0,015
17	-7,577	0,015
18	-7,465	0,015
19	-7,398	0,014

Continuação da Tabela B.3

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
20	-7,319	0,014
21	-7,318	0,015
22	-7,290	0,015
23	-7,247	0,015
24	-7,198	0,014
25	-7,211	0,015
26	-7,177	0,015
27	-7,100	0,014
28	-7,091	0,014
29	-7,049	0,013
30	-6,963	0,013
31	-6,945	0,012
32	-6,886	0,013
33	-6,803	0,012
34	-6,764	0,012
35	-6,668	0,011
36	-6,632	0,011
37	-6,579	0,011
38	-6,469	0,011
39	-6,412	0,009
40	-6,314	0,010
41	-6,324	0,008
42	-6,179	0,009
43	-6,165	0,008
44	-6,088	0,008
45	-5,988	0,008
46	-5,936	0,007
47	-5,870	0,007
48	-5,749	0,008

Continuação da Tabela B.3

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
49	-5,729	0,007
50	-5,596	0,008
51	-5,593	0,006
52	-5,456	0,007
53	-5,406	0,007
54	-5,351	0,007
55	-5,254	0,007
56	-5,176	0,007
57	-5,125	0,007
58	-5,006	0,007
59	-4,967	0,007
60	-4,775	0,009
61	-4,803	0,006
62	-4,647	0,008
63	-4,553	0,008
64	-4,449	0,008
65	-4,325	0,008
66	-4,242	0,008
67	-4,144	0,008
68	-4,002	0,008
69	-3,949	0,007
70	-3,715	0,009
71	-3,723	0,007
72	-3,520	0,008
73	-3,418	0,008
74	-3,278	0,008
75	-3,124	0,008
76	-3,030	0,007
77	-2,916	0,007

Continuação da Tabela B.3		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
78	-2,752	0,008
79	-2,694	0,006
80	-2,535	0,007
81	-2,548	0,004
82	-2,362	0,005
83	-2,274	0,005
84	-2,142	0,005
85	-2,000	0,005
86	-1,919	0,004
87	-1,831	0,003
88	-1,706	0,004
89	-1,661	0,002
90	-1,481	0,004
91	-1,540	0,000
92	-1,398	0,001
93	-1,343	0,001
94	-1,264	0,000
95	-1,128	0,002
96	-1,068	0,001
97	-0,999	0,001
98	-0,860	0,002
99	-0,886	0,000
100+	-0,773	0,000
Fim da Tabela		

Tabela B.4: Estimação do Parâmetro k_t do Método Lee-Carter (Sexo Feminino).

Início da Tabela	
Ano	\hat{k}_t
1940	106,529
1941	113,724
1942	113,920
1943	108,118
1944	103,923
1945	99,990
1946	98,677
1947	92,568
1948	87,491
1949	89,576
1950	79,164
1951	77,815
1952	68,997
1953	58,329
1954	55,963
1955	54,696
1956	55,969
1957	55,145
1958	42,263
1959	42,864
1960	37,435
1961	41,048
1962	36,956
1963	34,332
1964	35,494
1965	27,736
1966	28,284

Continuação da Tabela B.4

Ano	\hat{k}_t
1967	25,589
1968	21,322
1969	21,936
1970	19,159
1971	22,069
1972	11,928
1973	14,920
1974	10,898
1975	10,212
1976	12,390
1977	3,229
1978	1,816
1979	-1,551
1980	-3,194
1981	-4,961
1982	-8,385
1983	-10,050
1984	-11,096
1985	-16,108
1986	-16,401
1987	-17,331
1988	-21,674
1989	-23,011
1990	-20,824
1991	-20,356
1992	-26,653
1993	-25,881
1994	-28,789
1995	-31,409

Continuação da Tabela B.4

Ano	\hat{k}_t
1996	-30,547
1997	-32,038
1998	-32,542
1999	-38,676
2000	-45,347
2001	-49,389
2002	-48,622
2003	-51,611
2004	-58,144
2005	-63,988
2006	-67,048
2007	-71,637
2008	-73,302
2009	-74,634
2010	-75,423
2011	-83,683
2012	-83,440
2013	-88,883
2014	-95,702
2015	-96,514
2016	-99,127
2017	-98,894
2018	-88,798
2019	-94,626
2020	-91,794

Fim da Tabela

Tabela B.5: Estimação dos Parâmetros α_x e β_x do Método Lee-Carter (Sem Distinção de Sexo).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,835	0,026
1	-5,809	0,034
2	-6,440	0,031
3	-6,872	0,027
4	-7,138	0,026
5	-7,351	0,023
6	-7,511	0,023
7	-7,572	0,022
8	-7,701	0,020
9	-7,762	0,019
10	-7,778	0,018
11	-7,784	0,017
12	-7,801	0,018
13	-7,747	0,018
14	-7,575	0,017
15	-7,383	0,015
16	-7,189	0,014
17	-7,090	0,015
18	-6,908	0,013
19	-6,826	0,013
20	-6,774	0,013
21	-6,751	0,013
22	-6,739	0,013
23	-6,719	0,013
24	-6,685	0,013
25	-6,676	0,013
26	-6,650	0,013

Continuação da Tabela B.5

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
27	-6,614	0,013
28	-6,579	0,013
29	-6,569	0,012
30	-6,507	0,012
31	-6,498	0,011
32	-6,437	0,012
33	-6,374	0,011
34	-6,326	0,011
35	-6,255	0,011
36	-6,220	0,011
37	-6,165	0,010
38	-6,077	0,010
39	-6,023	0,009
40	-5,928	0,010
41	-5,915	0,008
42	-5,785	0,009
43	-5,747	0,008
44	-5,670	0,008
45	-5,579	0,008
46	-5,530	0,007
47	-5,458	0,007
48	-5,345	0,007
49	-5,300	0,007
50	-5,187	0,007
51	-5,179	0,006
52	-5,050	0,007
53	-4,981	0,006
54	-4,922	0,007
55	-4,832	0,007

Continuação da Tabela B.5

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
56	-4,760	0,007
57	-4,697	0,006
58	-4,596	0,007
59	-4,544	0,006
60	-4,389	0,008
61	-4,396	0,006
62	-4,255	0,007
63	-4,168	0,007
64	-4,083	0,007
65	-3,973	0,008
66	-3,896	0,008
67	-3,809	0,007
68	-3,685	0,008
69	-3,635	0,007
70	-3,438	0,009
71	-3,439	0,006
72	-3,262	0,008
73	-3,170	0,008
74	-3,052	0,008
75	-2,917	0,008
76	-2,833	0,007
77	-2,724	0,007
78	-2,582	0,008
79	-2,531	0,006
80	-2,397	0,008
81	-2,401	0,005
82	-2,234	0,006
83	-2,150	0,005
84	-2,028	0,005

Continuação da Tabela B.5		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
85	-1,905	0,005
86	-1,827	0,004
87	-1,746	0,004
88	-1,626	0,004
89	-1,585	0,002
90	-1,425	0,004
91	-1,477	0,000
92	-1,339	0,002
93	-1,287	0,001
94	-1,218	0,001
95	-1,096	0,002
96	-1,030	0,001
97	-0,971	0,001
98	-0,832	0,003
99	-0,857	0,000
100+	-0,742	0,000
Fim da Tabela		

Tabela B.6: Estimação do Parâmetro k_t do Método Lee-Carter (Sem Distinção de Sexo).

Início da Tabela	
Ano	\hat{k}_t
1940	88,552
1941	95,233
1942	94,719
1943	90,107
1944	86,934

Continuação da Tabela B.6

Ano	\hat{k}_t
1945	83,103
1946	81,722
1947	75,503
1948	71,710
1949	73,852
1950	64,653
1951	63,280
1952	55,150
1953	46,340
1954	44,269
1955	43,738
1956	45,217
1957	44,476
1958	33,166
1959	35,709
1960	30,136
1961	34,174
1962	31,082
1963	29,497
1964	30,298
1965	24,050
1966	25,414
1967	23,351
1968	19,555
1969	20,900
1970	18,061
1971	21,708
1972	12,267
1973	15,029

Continuação da Tabela B.6

Ano	\hat{k}_t
1974	12,066
1975	14,392
1976	14,450
1977	6,975
1978	5,927
1979	2,455
1980	1,993
1981	0,757
1982	-2,846
1983	-3,910
1984	-4,937
1985	-8,587
1986	-10,292
1987	-11,104
1988	-11,387
1989	-13,740
1990	-13,362
1991	-10,373
1992	-15,151
1993	-16,807
1994	-21,429
1995	-21,510
1996	-21,377
1997	-24,556
1998	-25,010
1999	-30,220
2000	-34,626
2001	-37,618
2002	-40,256

Continuação da Tabela B.6

Ano	\hat{k}_t
2003	-45,379
2004	-50,235
2005	-54,309
2006	-57,296
2007	-63,615
2008	-66,926
2009	-69,113
2010	-71,444
2011	-76,402
2012	-78,083
2013	-81,294
2014	-86,011
2015	-86,630
2016	-91,682
2017	-88,710
2018	-87,587
2019	-89,165
2020	-88,993

Fim da Tabela

Apêndice C

Estimação dos Parâmetros (Extensão do Método Lee-Carter ou Modelo de Poisson)

Tabela C.1: Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sexo Masculino).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,715	0,028
1	-6,111	0,047
2	-6,722	0,043
3	-7,058	0,036
4	-7,226	0,032
5	-7,386	0,028
6	-7,522	0,027
7	-7,499	0,024
8	-7,650	0,024
9	-7,664	0,022
10	-7,675	0,020
11	-7,661	0,019
12	-7,662	0,018
13	-7,555	0,017
14	-7,385	0,017

Continuação da Tabela C.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
15	-7,146	0,014
16	-6,867	0,012
17	-6,697	0,012
18	-6,505	0,011
19	-6,418	0,011
20	-6,379	0,011
21	-6,349	0,011
22	-6,341	0,011
23	-6,334	0,012
24	-6,311	0,012
25	-6,296	0,011
26	-6,275	0,011
27	-6,253	0,011
28	-6,210	0,011
29	-6,211	0,010
30	-6,156	0,010
31	-6,149	0,009
32	-6,089	0,010
33	-6,035	0,010
34	-5,988	0,009
35	-5,930	0,010
36	-5,889	0,010
37	-5,837	0,009
38	-5,764	0,010
39	-5,715	0,008
40	-5,620	0,009
41	-5,601	0,007
42	-5,475	0,008
43	-5,428	0,007

Continuação da Tabela C.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
44	-5,349	0,008
45	-5,262	0,008
46	-5,217	0,007
47	-5,142	0,006
48	-5,032	0,007
49	-4,974	0,006
50	-4,871	0,007
51	-4,859	0,005
52	-4,735	0,006
53	-4,654	0,006
54	-4,590	0,006
55	-4,503	0,006
56	-4,433	0,006
57	-4,360	0,006
58	-4,268	0,006
59	-4,204	0,006
60	-4,075	0,007
61	-4,057	0,005
62	-3,930	0,006
63	-3,843	0,007
64	-3,766	0,007
65	-3,665	0,007
66	-3,586	0,007
67	-3,501	0,007
68	-3,388	0,007
69	-3,328	0,006
70	-3,168	0,008
71	-3,145	0,006
72	-2,993	0,007

Continuação da Tabela C.1

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
73	-2,904	0,007
74	-2,799	0,007
75	-2,681	0,007
76	-2,596	0,007
77	-2,487	0,007
78	-2,364	0,007
79	-2,304	0,006
80	-2,201	0,007
81	-2,170	0,005
82	-2,033	0,006
83	-1,943	0,005
84	-1,833	0,005
85	-1,732	0,005
86	-1,653	0,004
87	-1,568	0,004
88	-1,457	0,004
89	-1,410	0,003
90	-1,294	0,004
91	-1,296	0,002
92	-1,174	0,002
93	-1,120	0,002
94	-1,076	0,001
95	-0,995	0,002
96	-0,907	0,001
97	-0,870	0,001
98	-0,717	0,002
99	-0,753	0,001

100+	-0,664	-0,001
Fim da Tabela		

Tabela C.2: Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sexo Masculino).

Início da Tabela	
Idade	\hat{k}_t
1940	67,356
1941	73,426
1942	70,375
1943	68,295
1944	66,251
1945	63,266
1946	63,022
1947	58,867
1948	58,044
1949	60,846
1950	54,533
1951	53,639
1952	53,200
1953	51,094
1954	46,742
1955	48,827
1956	48,612
1957	49,159
1958	43,796
1959	46,290
1960	41,565
1961	46,681

Continuação da Tabela C.2

Idade	\hat{k}_t
1962	41,742
1963	39,065
1964	38,065
1965	34,648
1966	35,612
1967	32,174
1968	31,206
1969	31,835
1970	30,224
1971	30,977
1972	20,713
1973	22,971
1974	20,373
1975	21,729
1976	19,985
1977	14,606
1978	11,481
1979	6,870
1980	7,079
1981	4,792
1982	0,097
1983	0,079
1984	-1,038
1985	-2,400
1986	-6,661
1987	-9,919
1988	-9,259
1989	-14,108
1990	-11,236

Continuação da Tabela C.2

Idade	\hat{k}_t
1991	-10,432
1992	-15,709
1993	-14,146
1994	-23,761
1995	-21,817
1996	-20,373
1997	-25,712
1998	-27,010
1999	-28,796
2000	-34,300
2001	-38,073
2002	-40,208
2003	-43,282
2004	-51,976
2005	-50,839
2006	-58,758
2007	-63,409
2008	-67,138
2009	-71,207
2010	-73,690
2011	-80,946
2012	-80,114
2013	-83,974
2014	-90,563
2015	-90,967
2016	-90,575
2017	-95,858
2018	-93,990
2019	-99,655

Continuação da Tabela C.2	
Idade	\hat{k}_t
2020	-88,313
Fim da Tabela	

Tabela C.3: Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sexo Feminino).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,862	0,020
1	-6,055	0,0318
2	-6,808	0,030
3	-7,205	0,026
4	-7,437	0,024
5	-7,630	0,022
6	-7,731	0,020
7	-7,867	0,019
8	-7,970	0,018
9	-8,073	0,018
10	-8,064	0,017
11	-8,063	0,016
12	-8,108	0,017
13	-8,048	0,017
14	-7,887	0,016
15	-7,763	0,016
16	-7,665	0,016
17	-7,661	0,017
18	-7,567	0,017
19	-7,483	0,017

Continuação da Tabela C.3

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
20	-7,422	0,017
21	-7,387	0,016
22	-7,390	0,017
23	-7,332	0,017
24	-7,258	0,016
25	-7,275	0,016
26	-7,217	0,016
27	-7,145	0,015
28	-7,138	0,015
29	-7,073	0,014
30	-6,990	0,014
31	-6,954	0,012
32	-6,902	0,013
33	-6,804	0,012
34	-6,762	0,011
35	-6,665	0,011
36	-6,628	0,011
37	-6,573	0,010
38	-6,464	0,010
39	-6,405	0,009
40	-6,307	0,009
41	-6,318	0,007
42	-6,171	0,008
43	-6,161	0,008
44	-6,081	0,007
45	-5,981	0,007
46	-5,933	0,007
47	-5,866	0,006
48	-5,747	0,007

Continuação da Tabela C.3

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
49	-5,729	0,006
50	-5,594	0,007
51	-5,591	0,005
52	-5,455	0,006
53	-5,405	0,006
54	-5,350	0,006
55	-5,254	0,006
56	-5,175	0,006
57	-5,124	0,006
58	-5,006	0,007
59	-4,965	0,006
60	-4,776	0,008
61	-4,799	0,006
62	-4,648	0,007
63	-4,552	0,007
64	-4,448	0,007
65	-4,326	0,007
66	-4,241	0,007
67	-4,142	0,007
68	-4,002	0,007
69	-3,939	0,006
70	-3,715	0,008
71	-3,710	0,006
72	-3,518	0,007
73	-3,413	0,007
74	-3,271	0,007
75	-3,121	0,007
76	-3,023	0,007
77	-2,908	0,006

Continuação da Tabela C.3

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
78	-2,748	0,007
79	-2,679	0,006
80	-2,534	0,006
81	-2,529	0,004
82	-2,356	0,005
83	-2,264	0,005
84	-2,134	0,005
85	-1,998	0,005
86	-1,911	0,004
87	-1,822	0,004
88	-1,702	0,004
89	-1,645	0,003
90	-1,480	0,003
91	-1,514	0,002
92	-1,387	0,002
93	-1,333	0,002
94	-1,258	0,001
95	-1,128	0,002
96	-1,068	0,001
97	-1,006	0,001
98	-0,867	0,002
99	-0,900	0,001
100+	-0,780	0,000

Fim da Tabela

Tabela C.4: Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sexo Feminino).

Início da Tabela	
Idade	\hat{k}_t
1940	98,464
1941	107,544
1942	103,243
1943	99,473
1944	95,138
1945	91,436
1946	91,606
1947	86,004
1948	83,780
1949	87,827
1950	78,496
1951	77,551
1952	75,829
1953	70,992
1954	65,608
1955	69,405
1956	69,670
1957	69,237
1958	60,021
1959	63,378
1960	57,110
1961	63,977
1962	56,278
1963	53,243
1964	52,163
1965	46,427
1966	47,889

Continuação da Tabela C.4

Idade	\hat{k}_t
1967	42,917
1968	41,911
1969	43,731
1970	40,422
1971	40,735
1972	27,813
1973	31,899
1974	27,524
1975	25,548
1976	24,713
1977	15,509
1978	14,336
1979	8,254
1980	6,262
1981	3,798
1982	-3,096
1983	-1,797
1984	-5,110
1985	-7,956
1986	-11,421
1987	-15,982
1988	-17,167
1989	-23,603
1990	-17,384
1991	-18,973
1992	-28,015
1993	-24,967
1994	-36,992
1995	-34,920

Continuação da Tabela C.4

Idade	\hat{k}_t
1996	-35,582
1997	-40,105
1998	-42,515
1999	-44,005
2000	-51,233
2001	-56,247
2002	-58,666
2003	-58,095
2004	-73,992
2005	-70,355
2006	-86,540
2007	-86,806
2008	-90,940
2009	-94,136
2010	-98,228
2011	-107,401
2012	-105,256
2013	-112,688
2014	-120,606
2015	-118,115
2016	-120,332
2017	-125,767
2018	-125,156
2019	-130,427
2020	-116,581

Fim da Tabela

Tabela C.5: Estimação dos Parâmetros α_x e β_x da Extensão do Método Lee-Carter (Sem Distinção de Sexo).

Início da Tabela		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
0	-3,777	0,025
1	-6,075	0,040
2	-6,755	0,037
3	-7,124	0,032
4	-7,319	0,029
5	-7,493	0,029
6	-7,616	0,024
7	-7,657	0,022
8	-7,789	0,021
9	-7,837	0,021
10	-7,842	0,019
11	-7,833	0,018
12	-7,849	0,018
13	-7,768	0,018
14	-7,592	0,017
15	-7,389	0,015
16	-7,163	0,014
17	-7,035	0,013
18	-6,864	0,013
19	-6,783	0,012
20	-6,740	0,013
21	-6,713	0,013
22	-6,705	0,013
23	-6,690	0,013
24	-6,657	0,013
25	-6,648	0,013
26	-6,617	0,012

Continuação da Tabela C.5

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
27	-6,585	0,012
28	-6,553	0,012
29	-6,541	0,011
30	-6,477	0,012
31	-6,469	0,010
32	-6,411	0,011
33	-6,346	0,011
34	-6,302	0,010
35	-6,234	0,010
36	-6,195	0,010
37	-6,145	0,010
38	-6,061	0,010
39	-6,010	0,009
40	-5,914	0,009
41	-5,908	0,007
42	-5,775	0,008
43	-5,741	0,007
44	-5,663	0,007
45	-5,572	0,008
46	-5,527	0,007
47	-5,456	0,006
48	-5,342	0,007
49	-5,300	0,006
50	-5,186	0,007
51	-5,178	0,005
52	-5,049	0,006
53	-4,981	0,006
54	-4,921	0,006
55	-4,832	0,006

Continuação da Tabela C.5

Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
56	-4,759	0,006
57	-4,696	0,006
58	-4,595	0,006
59	-4,542	0,006
60	-4,389	0,007
61	-4,391	0,005
62	-4,255	0,007
63	-4,166	0,007
64	-4,081	0,007
65	-3,972	0,007
66	-3,894	0,007
67	-3,805	0,007
68	-3,683	0,007
69	-3,683	0,006
70	-3,438	0,008
71	-3,428	0,006
72	-3,260	0,007
73	-3,166	0,007
74	-3,047	0,007
75	-2,915	0,008
76	-2,827	0,007
77	-2,719	0,007
78	-2,579	0,007
79	-2,519	0,006
80	-2,396	0,007
81	-2,384	0,005
82	-2,229	0,006
83	-2,142	0,005
84	-2,023	0,005

Continuação da Tabela C.5		
Idade	$\hat{\alpha}_x$	$\hat{\beta}_x$
85	-1,904	0,005
86	-1,823	0,004
87	-1,739	0,004
88	-1,624	0,004
89	-1,573	0,003
90	-1,426	0,004
91	-1,452	0,002
92	-1,330	0,002
93	-1,278	0,002
94	-1,213	0,001
95	-1,097	0,002
96	-1,032	0,002
97	-0,978	0,001
98	-0,837	0,002
99	-0,871	0,001
100+	-0,760	0,000
Fim da Tabela		

Tabela C.6: Estimação do Parâmetro k_t da Extensão do Método Lee-Carter (Sem Distinção de Sexo).

Início da Tabela	
Idade	\hat{k}_t
1940	78,041
1941	85,288
1942	81,696
1943	79,005
1944	76,129

Continuação da Tabela C.6

Idade	\hat{k}_t
1945	72,877
1946	72,781
1947	68,042
1948	66,762
1949	70,076
1950	62,626
1951	61,722
1952	60,881
1953	57,753
1954	53,010
1955	55,841
1956	55,754
1957	55,962
1958	49,148
1959	52,058
1960	46,693
1961	52,568
1962	46,530
1963	43,760
1964	42,712
1965	38,407
1966	39,575
1967	35,512
1968	34,548
1969	35,628
1970	33,404
1971	33,948
1972	22,578
1973	25,637

Continuação da Tabela C.6

Idade	\hat{k}_t
1974	22,363
1975	22,612
1976	21,197
1977	14,282
1978	11,977
1979	6,801
1980	6,170
1981	4,031
1982	-1,594
1983	-1,010
1984	-2,988
1985	-4,964
1986	-8,756
1987	-12,358
1988	-12,339
1989	-17,738
1990	-13,492
1991	-13,523
1992	-20,180
1993	-18,096
1994	-28,416
1995	-26,367
1996	-25,732
1997	-30,691
1998	-32,374
1999	-33,918
2000	-40,014
2001	-44,172
2002	-46,448

Continuação da Tabela C.6

Idade	\hat{k}_t
2003	-47,970
2004	-59,467
2005	-57,234
2006	-68,321
2007	-70,998
2008	-74,820
2009	-78,471
2010	-81,393
2011	-89,329
2012	-87,706
2013	-92,981
2014	-99,921
2015	-99,066
2016	-99,680
2017	-104,979
2018	-103,490
2019	-108,826
2020	-96,561

Fim da Tabela

Apêndice D

Tábuas de Mortalidade

Tabela D.1: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Masculino.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,05715	0,0545	0,00001	5450,348	95367,204	6342711,384	63,427
1	0,00926	0,00922	94549,652	871,569	94113,868	6247344,179	66,075
2	0,00435	0,00434	93678,083	406,721	93474,722	6153230,312	65,685
3	0,00257	0,00257	93271,362	239,749	93151,487	6059755,589	64,969
4	0,00189	0,00189	93031,613	175,879	92943,674	5966604,102	64,135
5	0,00145	0,00145	92855,734	134,684	92788,393	5873660,428	63,256
6	0,00122	0,00122	92721,051	113,192	92664,455	5780872,036	62,347
7	0,00115	0,00115	92607,858	106,128	92554,795	5688207,581	61,423
8	0,00097	0,00097	92501,731	89,763	92456,849	5595652,787	60,492
9	0,00091	0,00091	92411,968	83,683	92370,126	5503195,937	59,551
10	0,00086	0,00085	92328,284	78,915	92288,827	5410825,812	58,604
11	0,00084	0,00084	92249,37	77,581	92210,579	5318536,985	57,654
12	0,00082	0,00082	92171,789	75,502	92134,038	5226326,405	56,702
13	0,00088	0,00088	92096,286	80,987	92055,793	5134192,368	55,748
14	0,00102	0,00102	92015,299	94,101	91968,249	5042136,575	54,797
15	0,0012	0,00119	91921,198	109,813	91866,291	4950168,326	53,852

Continuação da Tabela D.1

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
16	0,0015	0,0015	91811,384	137,688	91742,54	4858302,035	52,916
17	0,00175	0,00175	91673,696	160,591	91593,4	4766559,495	51,995
18	0,00206	0,00206	91513,105	188,482	91418,864	4674966,095	51,085
19	0,00224	0,00224	91324,622	204,711	91222,267	4583547,231	50,19
20	0,00235	0,00235	91119,911	214,065	91012,878	4492324,964	49,301
21	0,00244	0,00243	90905,846	221,22	90795,235	4401312,086	48,416
22	0,00249	0,00248	90684,625	225,338	90571,957	4310516,851	47,533
23	0,00252	0,00252	90459,288	227,941	90345,317	4219944,894	46,65
24	0,00259	0,00259	90231,347	233,474	90114,61	4129599,577	45,767
25	0,00257	0,00257	89997,873	230,88	89882,433	4039484,967	44,884
26	0,0026	0,0026	89766,993	232,967	89650,51	3949602,534	43,998
27	0,00264	0,00264	89534,026	236,348	89415,852	3859952,024	43,112
28	0,00278	0,00278	89297,678	248,072	89173,642	3770536,172	42,224
29	0,00271	0,00271	89049,606	241,07	88929,071	3681362,53	41,341
30	0,00288	0,00288	88808,536	255,495	88680,788	3592433,459	40,451
31	0,00284	0,00283	88553,041	250,963	88427,559	3503752,671	39,567
32	0,00309	0,00309	88302,078	272,715	88165,721	3415325,112	38,678
33	0,00325	0,00325	88029,363	285,75	87886,488	3327159,391	37,796
34	0,00332	0,00332	87743,613	290,912	87598,157	3239272,903	36,917
35	0,00356	0,00355	87452,701	310,481	87297,46	3151674,746	36,039
36	0,00371	0,0037	87142,22	322,514	86980,963	3064377,285	35,165
37	0,00387	0,00386	86819,705	335,476	86651,968	2977396,323	34,294
38	0,0042	0,00419	86484,23	362,05	86303,205	2890744,355	33,425
39	0,00425	0,00424	86122,18	364,863	85939,748	2804441,15	32,564
40	0,00474	0,00473	85757,317	405,69	85554,472	2718501,402	31,7
41	0,0046	0,00459	85351,627	391,618	85155,818	2632946,93	30,848
42	0,00541	0,0054	84960,009	458,652	84730,683	2547791,111	29,988
43	0,00548	0,00547	84501,358	461,968	84270,373	2463060,428	29,148
44	0,00597	0,00596	84039,389	500,464	83789,157	2378790,055	28,306

Continuação da Tabela D.1

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
45	0,00655	0,00652	83538,926	545,004	83266,423	2295000,897	27,472
46	0,00662	0,00659	82993,921	547,343	82720,25	2211734,474	26,649
47	0,00708	0,00706	82446,578	582,067	82155,545	2129014,224	25,823
48	0,00799	0,00796	81864,512	651,878	81538,573	2046858,679	25,003
49	0,00831	0,00827	81212,634	671,972	80876,648	1965320,106	24,2
50	0,0094	0,00935	80540,662	753,457	80163,934	1884443,458	23,397
51	0,00903	0,00899	79787,205	717,289	79428,561	1804279,524	22,614
52	0,01055	0,01049	79069,916	829,43	78655,201	1724850,963	21,814
53	0,01129	0,01123	78240,486	878,708	77801,132	1646195,762	21,04
54	0,01215	0,01208	77361,778	934,579	76894,488	1568394,63	20,274
55	0,01326	0,01317	76427,199	1006,873	75923,762	1491500,142	19,515
56	0,0143	0,0142	75420,325	1070,949	74884,851	1415576,38	18,769
57	0,01515	0,01504	74349,377	1118,159	73790,297	1340691,529	18,032
58	0,01685	0,01671	73231,218	1223,353	72619,541	1266901,231	17,3
59	0,01775	0,0176	72007,865	1267,188	71374,271	1194281,69	16,585
60	0,02091	0,0207	70740,677	1464,033	70008,66	1122907,419	15,874
61	0,02025	0,02005	69276,644	1388,991	68582,148	1052898,759	15,198
62	0,02382	0,02354	67887,653	1597,892	67088,707	984316,61	14,499
63	0,02609	0,02576	66289,761	1707,459	65436,031	917227,904	13,837
64	0,02828	0,02789	64582,302	1800,902	63681,851	851791,872	13,189
65	0,03144	0,03095	62781,401	1943,063	61809,869	788110,021	12,553
66	0,03409	0,03352	60838,337	2039,397	59818,639	726300,152	11,938
67	0,03682	0,03615	58798,94	2125,584	57736,148	666481,513	11,335
68	0,04188	0,04102	56673,356	2324,561	55511,076	608745,365	10,741
69	0,04347	0,04255	54348,795	2312,326	53192,632	553234,289	10,179
70	0,05316	0,05178	52036,47	2694,629	50689,155	500041,657	9,609
71	0,05191	0,0506	49341,841	2496,541	48093,571	449352,502	9,107
72	0,06269	0,06079	46845,3	2847,531	45421,535	401258,931	8,566
73	0,06781	0,06558	43997,769	2885,547	42554,996	355837,396	8,088

Continuação da Tabela D.1

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
74	0,07541	0,07267	41112,222	2987,701	39618,372	313282,401	7,62
75	0,08544	0,08194	38124,521	3123,886	36562,578	273664,029	7,178
76	0,09185	0,08781	35000,635	3073,52	33463,875	237101,451	6,774
77	0,10258	0,09758	31927,114	3115,388	30369,42	203637,577	6,378
78	0,11701	0,11054	28811,726	3184,845	27219,304	173268,157	6,014
79	0,12037	0,11353	25626,881	2909,527	24172,118	146048,853	5,699
80	0,13539	0,1268	22717,354	2880,647	21277,03	121876,735	5,365
81	0,13338	0,12504	19836,707	2480,416	18596,499	100599,705	5,071
82	0,15522	0,14404	17356,291	2499,969	16106,307	82003,206	4,725
83	0,1676	0,15464	14856,322	2297,433	13707,606	65896,9	4,436
84	0,1877	0,1716	12558,889	2155,095	11481,342	52189,294	4,156
85	0,20689	0,18749	10403,794	1950,627	9428,481	40707,952	3,913
86	0,21883	0,19724	8453,167	1667,34	7619,497	31279,471	3,7
87	0,23761	0,21238	6785,827	1441,147	6065,253	23659,974	3,487
88	0,26598	0,23476	5344,68	1254,7	4717,33	17594,721	3,292
89	0,26885	0,23699	4089,98	969,279	3605,34	12877,391	3,149
90	0,30542	0,26496	3120,701	826,862	2707,27	9272,051	2,971
91	0,29117	0,25417	2293,839	583,018	2002,33	6564,781	2,862
92	0,33346	0,2858	1710,821	488,96	1466,341	4562,451	2,667
93	0,34663	0,29543	1221,861	360,972	1041,375	3096,109	2,534
94	0,35402	0,30078	860,889	258,94	731,419	2054,734	2,387
95	0,38699	0,32425	601,949	195,182	504,358	1323,315	2,198
96	0,4223	0,34868	406,767	141,831	335,852	818,957	2,013
97	0,43124	0,35475	264,936	93,985	217,944	483,105	1,823
98	0,52371	0,41503	170,951	70,95	135,476	265,162	1,551
99	0,4787	0,38625	100,001	38,626	80,688	129,686	1,297
100	0,50525	0,40335	61,375	24,756	48,997	0	0

Fim da Tabela

Tabela D.2: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Masculino.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,03962	0,03833	0,00001	3832,888	96742,045	6602040,677	66,02
1	0,0062	0,00618	96167,112	594,738	95869,743	6505298,631	67,646
2	0,00314	0,00314	95572,373	299,653	95422,547	6409428,889	67,064
3	0,00192	0,00192	95272,72	182,594	95181,423	6314006,342	66,273
4	0,00143	0,00143	95090,127	135,85	95022,201	6218824,918	65,399
5	0,00112	0,00111	94954,276	105,833	94901,36	6123802,717	64,492
6	0,00093	0,00093	94848,443	88,182	94804,352	6028901,357	63,564
7	0,00089	0,00089	94760,262	84,592	94717,966	5934097,005	62,622
8	0,00076	0,00076	94675,67	71,732	94639,803	5839379,039	61,678
9	0,00071	0,00071	94603,937	67,439	94570,218	5744739,236	60,724
10	0,00068	0,00068	94536,498	64,224	94504,386	5650169,018	59,767
11	0,00068	0,00068	94472,274	64,035	94440,256	5555664,632	58,807
12	0,00066	0,00066	94408,239	62,433	94377,023	5461224,376	57,847
13	0,00071	0,00071	94345,806	67,087	94312,263	5366847,353	56,885
14	0,00084	0,00084	94278,72	79,215	94239,112	5272535,09	55,925
15	0,001	0,001	94199,505	94,485	94152,262	5178295,978	54,972
16	0,00126	0,00126	94105,02	118,381	94045,829	5084143,716	54,026
17	0,00146	0,00146	93986,638	137,476	93917,9	4990097,887	53,094
18	0,00173	0,00173	93849,162	162,346	93767,989	4896179,987	52,171
19	0,00189	0,00189	93686,816	176,923	93598,355	4802411,997	51,26
20	0,00198	0,00198	93509,893	184,814	93417,487	4708813,643	50,356
21	0,00204	0,00203	93325,08	189,801	93230,179	4615396,156	49,455
22	0,00205	0,00205	93135,278	190,72	93039,918	4522165,977	48,555
23	0,00208	0,00208	92944,558	193,542	92847,788	4429126,058	47,653
24	0,00217	0,00217	92751,017	201,363	92650,335	4336278,271	46,752

Continuação da Tabela D.2

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	\dot{e}_{1970}
25	0,00217	0,00217	92549,653	200,913	92449,197	4243627,936	45,852
26	0,00221	0,00221	92348,741	204,194	92246,644	4151178,739	44,951
27	0,00225	0,00225	92144,547	207,366	92040,864	4058932,095	44,05
28	0,00237	0,00237	91937,181	218,028	91828,167	3966891,231	43,148
29	0,00234	0,00234	91719,153	214,48	91611,913	3875063,064	42,249
30	0,00249	0,00248	91504,673	227,125	91391,11	3783451,151	41,347
31	0,00248	0,00248	91277,548	226,081	91164,507	3692060,041	40,449
32	0,00268	0,00267	91051,467	243,479	90929,727	3600895,533	39,548
33	0,00282	0,00281	90807,988	255,431	90680,272	3509965,806	38,653
34	0,00292	0,00291	90552,557	263,828	90420,643	3419285,534	37,76
35	0,00311	0,00311	90288,729	280,558	90148,45	3328864,891	36,869
36	0,00325	0,00324	90008,171	291,725	89862,308	3238716,441	35,982
37	0,00342	0,00341	89716,446	306,062	89563,415	3148854,132	35,098
38	0,0037	0,00369	89410,383	329,848	89245,459	3059290,718	34,216
39	0,00382	0,00381	89080,535	339,553	88910,759	2970045,259	33,341
40	0,00423	0,00422	88740,982	374,603	88553,681	2881134,5	32,467
41	0,00422	0,00421	88366,38	372,294	88180,233	2792580,819	31,602
42	0,00488	0,00487	87994,086	428,129	87780,021	2704400,586	30,734
43	0,00503	0,00502	87565,956	439,437	87346,238	2616620,565	29,882
44	0,00546	0,00545	87126,52	474,72	86889,16	2529274,327	29,03
45	0,00598	0,00596	86651,799	516,445	86393,577	2442385,168	28,186
46	0,00615	0,00613	86135,355	528,034	85871,338	2355991,591	27,352
47	0,00661	0,00659	85607,321	564,205	85325,218	2270120,253	26,518
48	0,00742	0,00739	85043,116	628,745	84728,743	2184795,034	25,69
49	0,00778	0,00775	84414,37	654,439	84087,151	2100066,291	24,878
50	0,00874	0,0087	83759,931	728,939	83395,462	2015979,14	24,069
51	0,00855	0,00851	83030,993	706,838	82677,573	1932583,678	23,275
52	0,00989	0,00984	82324,154	810,088	81919,11	1849906,105	22,471
53	0,01064	0,01058	81514,066	862,649	81082,741	1767986,995	21,689

Continuação da Tabela D.2

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	\dot{e}_{1970}
54	0,01139	0,01132	80651,417	913,103	80194,865	1686904,254	20,916
55	0,01244	0,01236	79738,313	985,628	79245,499	1606709,389	20,15
56	0,01338	0,01329	78752,686	1046,741	78229,315	1527463,889	19,396
57	0,01423	0,01412	77705,944	1097,589	77157,149	1449234,574	18,65
58	0,01576	0,01564	76608,355	1198,136	76009,287	1372077,425	17,91
59	0,0166	0,01646	75410,219	1241,443	74789,497	1296068,138	17,187
60	0,01942	0,01923	74168,776	1426,24	73455,655	1221278,641	16,466
61	0,01893	0,01875	72742,535	1364,193	72060,439	1147822,985	15,779
62	0,02217	0,02192	71378,342	1564,793	70595,946	1075762,547	15,071
63	0,02421	0,02393	69813,549	1670,297	68978,401	1005166,601	14,398
64	0,02621	0,02587	68143,252	1762,951	67261,776	936188,2	13,739
65	0,0291	0,02869	66380,3	1904,186	65428,207	868926,424	13,09
66	0,03151	0,03102	64476,114	1999,869	63476,179	803498,217	12,462
67	0,03401	0,03344	62476,245	2089,309	61431,59	740022,038	11,845
68	0,03858	0,03785	60386,936	2285,879	59243,997	678590,447	11,237
69	0,03994	0,03916	58101,057	2275,323	56963,396	619346,45	10,66
70	0,04877	0,0476	55825,734	2657,555	54496,956	562383,055	10,074
71	0,04761	0,0465	53168,179	2472,451	51931,953	507886,098	9,552
72	0,05763	0,05601	50695,728	2839,587	49275,934	455954,145	8,994
73	0,06233	0,06045	47856,141	2892,694	46409,794	406678,211	8,498
74	0,06922	0,0669	44963,447	3008,209	43459,342	360268,417	8,012
75	0,07867	0,07569	41955,237	3175,544	40367,465	316809,075	7,551
76	0,08421	0,08081	38779,693	3133,818	37212,784	276441,61	7,129
77	0,09425	0,09000	35645,875	3208,306	34041,722	239228,826	6,711
78	0,10765	0,10215	32437,569	3313,465	30780,837	205187,104	6,326
79	0,10983	0,10411	29124,104	3032,148	27608,03	174406,267	5,988
80	0,12587	0,11842	26091,956	3089,784	24547,064	146798,237	5,626
81	0,12245	0,11538	23002,172	2654,075	21675,134	122251,173	5,315
82	0,14472	0,13495	20348,096	2746,057	18975,068	100576,039	4,943

Continuação da Tabela D.2

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	\dot{e}_{1970}
83	0,15612	0,14482	17602,039	2549,09	16327,494	81600,971	4,636
84	0,17623	0,16196	15052,949	2437,901	13833,999	65273,477	4,336
85	0,19544	0,17804	12615,048	2245,984	11492,056	51439,478	4,078
86	0,20737	0,18789	10369,064	1948,248	9394,94	39947,422	3,853
87	0,22373	0,20122	8420,816	1694,46	7573,586	30552,482	3,628
88	0,25388	0,22528	6726,356	1515,327	5968,693	22978,896	3,416
89	0,25573	0,22674	5211,029	1181,554	4620,252	17010,204	3,264
90	0,29647	0,25819	4029,475	1040,389	3509,28	12389,952	3,075
91	0,27183	0,2393	2989,086	715,294	2631,439	8880,671	2,971
92	0,31902	0,27514	2273,792	625,604	1960,99	6249,232	2,748
93	0,33115	0,28411	1648,189	468,261	1414,058	4288,242	2,602
94	0,34446	0,29385	1179,928	346,726	1006,565	2874,183	2,436
95	0,3789	0,31855	833,202	265,415	700,495	1867,619	2,241
96	0,41358	0,34271	567,787	194,585	470,495	1167,124	2,056
97	0,41844	0,34604	373,202	129,144	308,63	696,629	1,867
98	0,49604	0,39746	244,059	97,004	195,556	387,999	1,59
99	0,47018	0,38069	147,054	55,982	119,063	192,443	1,309
100	0,48223	0,38855	91,072	35,386	73,379	0	0
Fim da Tabela							

Tabela D.3: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Feminino.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,04668	0,04492	0,00001	4491,955	96226,758	7000690,988	70,007
1	0,00848	0,00845	95508,045	806,695	95104,698	6904464,23	72,292
2	0,00373	0,00372	94701,35	352,51	94525,095	6809359,533	71,904

Continuação da Tabela D.3

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
3	0,00216	0,00216	94348,84	203,724	94246,978	6714834,437	71,17
4	0,00155	0,00155	94145,116	146,191	94072,02	6620587,459	70,323
5	0,00116	0,00116	93998,925	108,814	93944,518	6526515,439	69,432
6	0,00098	0,00098	93890,111	92,357	93843,933	6432570,921	68,512
7	0,00084	0,00084	93797,754	78,972	93758,268	6338726,989	67,579
8	0,00072	0,00072	93718,782	67,914	93684,825	6244968,721	66,635
9	0,00066	0,00066	93650,867	61,525	93620,105	6151283,896	65,683
10	0,00062	0,00062	93589,343	57,947	93560,369	6057663,791	64,726
11	6e-04	6e-04	93531,396	56,437	93503,177	5964103,422	63,766
12	6e-04	6e-04	93474,959	56,253	93446,833	5870600,245	62,804
13	0,00064	0,00064	93418,706	59,731	93388,841	5777153,412	61,842
14	0,00072	0,00072	93358,976	67,295	93325,329	5683764,571	60,881
15	8e-04	8e-04	93291,681	75,027	93254,168	5590439,242	59,924
16	9e-04	9e-04	93216,654	83,58	93174,865	5497185,074	58,972
17	0,00094	0,00094	93133,075	87,369	93089,39	5404010,21	58,025
18	0,00104	0,00104	93045,706	97,121	92997,145	5310920,82	57,079
19	0,0011	0,0011	92948,584	102,432	92897,368	5217923,675	56,138
20	0,00117	0,00117	92846,153	108,876	92791,714	5125026,307	55,199
21	0,0012	0,0012	92737,276	111,07	92681,741	5032234,592	54,263
22	0,00125	0,00125	92626,206	115,594	92568,409	4939552,851	53,328
23	0,00128	0,00128	92510,612	118,353	92451,435	4846984,442	52,394
24	0,00134	0,00134	92392,259	124,148	92330,185	4754533,006	51,46
25	0,00133	0,00133	92268,111	122,356	92206,933	4662202,822	50,529
26	0,00139	0,00139	92145,755	127,853	92081,828	4569995,889	49,595
27	0,00145	0,00145	92017,901	133,494	91951,155	4477914,061	48,664
28	0,00147	0,00147	91884,408	135,218	91816,799	4385962,906	47,733
29	0,00146	0,00146	91749,19	134,164	91682,108	4294146,108	46,803
30	0,00161	0,00161	91615,026	147,377	91541,338	4202464	45,871
31	0,00153	0,00153	91467,649	140,229	91397,535	4110922,662	44,944

Continuação da Tabela D.3

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
32	0,00168	0,00168	91327,42	153,522	91250,659	4019525,128	44,012
33	0,0018	0,0018	91173,898	163,996	91091,9	3928274,469	43,086
34	0,00183	0,00183	91009,902	166,274	90926,765	3837182,569	42,162
35	0,00197	0,00197	90843,628	178,975	90754,14	3746255,804	41,239
36	0,00205	0,00204	90664,653	185,267	90572,019	3655501,664	40,319
37	0,00211	0,00211	90479,386	190,773	90383,999	3564929,645	39,4
38	0,00234	0,00234	90288,613	211,371	90182,927	3474545,646	38,483
39	0,00235	0,00235	90077,242	211,651	89971,416	3384362,718	37,572
40	0,00265	0,00264	89865,591	237,534	89746,824	3294391,302	36,659
41	0,00243	0,00242	89628,057	217,242	89519,436	3204644,478	35,755
42	0,00291	0,00291	89410,815	260,252	89280,689	3115125,043	34,841
43	0,00286	0,00286	89150,563	254,913	89023,106	3025844,354	33,941
44	0,00306	0,00305	88895,649	271,463	88759,918	2936821,248	33,037
45	0,0034	0,00339	88624,186	300,835	88473,769	2848061,33	32,136
46	0,00346	0,00346	88323,351	305,217	88170,742	2759587,562	31,244
47	0,00362	0,00361	88018,134	317,655	87859,306	2671416,819	30,351
48	0,00418	0,00417	87700,479	366,01	87517,474	2583557,513	29,459
49	0,00416	0,00415	87334,469	362,504	87153,217	2496040,038	28,58
50	0,00491	0,0049	86971,965	426,353	86758,789	2408886,821	27,697
51	0,00465	0,00463	86545,613	401,115	86345,055	2322128,032	26,831
52	0,00555	0,00554	86144,498	477,039	85905,978	2235782,977	25,954
53	0,00576	0,00574	85667,459	491,862	85421,527	2149876,999	25,096
54	0,00612	0,0061	85175,596	519,396	84915,898	2064455,471	24,238
55	0,00676	0,00674	84656,201	570,526	84370,937	1979539,573	23,383
56	0,00734	0,00731	84085,674	614,829	83778,26	1895168,635	22,539
57	0,00761	0,00758	83470,846	632,437	83154,627	1811390,375	21,701
58	0,00873	0,00869	82838,409	720,143	82478,337	1728235,748	20,863
59	0,00885	0,00881	82118,266	723,591	81756,47	1645757,411	20,041
60	0,01152	0,01145	81394,675	931,942	80928,704	1564000,941	19,215

Continuação da Tabela D.3

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
61	0,01031	0,01025	80462,733	824,96	80050,253	1483072,237	18,432
62	0,0127	0,01262	79637,772	1005,373	79135,086	1403021,984	17,618
63	0,01394	0,01385	78632,399	1088,799	78088	1323886,898	16,836
64	0,01544	0,01532	77543,6	1187,833	76949,684	1245798,899	16,066
65	0,01784	0,01768	76355,768	1350,213	75680,661	1168849,215	15,308
66	0,01916	0,01898	75005,555	1423,785	74293,662	1093168,553	14,575
67	0,02097	0,02076	73581,77	1527,285	72818,127	1018874,891	13,847
68	0,02459	0,02429	72054,484	1750,464	71179,252	946056,764	13,13
69	0,02506	0,02475	70304,02	1740,238	69433,901	874877,511	12,444
70	0,03404	0,03347	68563,782	2294,867	67416,349	805443,61	11,747
71	0,03123	0,03075	66268,915	2037,825	65250,003	738027,261	11,137
72	0,03973	0,03895	64231,091	2502,121	62980,03	672777,258	10,474
73	0,04352	0,0426	61728,969	2629,395	60414,272	609797,228	9,879
74	0,04993	0,04871	59099,574	2878,917	57660,115	549382,957	9,296
75	0,05943	0,05772	56220,657	3244,859	54598,227	491722,841	8,746
76	0,06364	0,06168	52975,798	3267,44	51342,078	437124,614	8,251
77	0,07079	0,06837	49708,358	3398,712	48009,001	385782,536	7,761
78	0,08447	0,08105	46309,645	3753,25	44433,02	337773,535	7,294
79	0,08619	0,08263	42556,395	3516,361	40798,215	293340,515	6,893
80	0,10313	0,09808	39040,034	3828,928	37125,571	252542,3	6,469
81	0,09558	0,09122	35211,107	3211,915	33605,149	215416,729	6,118
82	0,11606	0,1097	31999,192	3510,262	30244,061	181811,58	5,682
83	0,12479	0,11746	28488,93	3346,396	26815,732	151567,519	5,32
84	0,14229	0,13284	25142,534	3339,96	23472,554	124751,787	4,962
85	0,16342	0,15108	21802,575	3293,867	20155,641	101279,233	4,645
86	0,17418	0,16023	18508,708	2965,582	17025,917	81123,592	4,383
87	0,18685	0,17088	15543,126	2656,074	14215,089	64097,675	4,124
88	0,21149	0,19127	12887,052	2464,883	11654,61	49882,586	3,871
89	0,2157	0,1947	10422,168	2029,223	9407,557	38227,976	3,668

Continuação da Tabela D.3

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
90	0,26098	0,23086	8392,945	1937,586	7424,153	28820,419	3,434
91	0,23553	0,21072	6455,36	1360,266	5775,227	21396,267	3,314
92	0,27146	0,23902	5095,094	1217,827	4486,18	15621,04	3,066
93	0,28096	0,24635	3877,267	955,171	3399,681	11134,86	2,872
94	0,30004	0,2609	2922,095	762,384	2540,903	7735,179	2,647
95	0,34792	0,29637	2159,711	640,068	1839,677	5194,276	2,405
96	0,36487	0,30858	1519,643	468,928	1285,179	3354,598	2,207
97	0,3844	0,32243	1050,715	338,784	881,322	2069,42	1,97
98	0,45446	0,37032	711,93	263,639	580,111	1188,097	1,669
99	0,42089	0,34772	448,291	155,878	370,352	607,987	1,356
100	0,46103	0,37466	292,413	109,556	237,635	0	0

Fim da Tabela

Tabela D.4: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Feminino.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,03027	0,02952	0,00001	2952,115	97520,223	7312296,503	73,123
1	0,00503	0,00502	97047,885	487,153	96804,308	7214776,28	74,342
2	0,0024	0,0024	96560,731	231,94	96444,761	7117971,972	73,715
3	0,00149	0,00149	96328,792	143,633	96256,975	7021527,211	72,891
4	0,00108	0,00108	96185,158	103,554	96133,381	6925270,236	71,999
5	0,00082	0,00082	96081,604	79,214	96041,997	6829136,854	71,076
6	0,00073	0,00073	96002,39	69,644	95967,569	6733094,857	70,135
7	0,00061	0,00061	95932,747	58,962	95903,266	6637127,289	69,185
8	0,00054	0,00054	95873,785	51,656	95847,957	6541224,023	68,227
9	0,00049	0,00049	95822,129	47,154	95798,552	6445376,066	67,264

Continuação da Tabela D.4

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
10	0,00047	0,00047	95774,975	45,426	95752,262	6349577,514	66,297
11	0,00046	0,00046	95729,549	43,631	95707,733	6253825,252	65,328
12	0,00045	0,00045	95685,917	43,249	95664,293	6158117,519	64,358
13	0,00048	0,00048	95642,669	45,525	95619,906	6062453,226	63,386
14	0,00054	0,00053	95597,143	51,134	95571,577	5966833,319	62,416
15	0,00061	0,00061	95546,01	57,837	95517,091	5871261,743	61,45
16	0,00066	0,00066	95488,173	62,87	95456,738	5775744,651	60,486
17	0,00069	0,00069	95425,303	65,508	95392,549	5680287,913	59,526
18	0,00076	0,00076	95359,795	72,611	95323,489	5584895,365	58,567
19	0,00081	0,00081	95287,184	76,882	95248,743	5489571,875	57,611
20	0,00086	0,00086	95210,302	82,012	95169,296	5394323,132	56,657
21	0,00088	0,00088	95128,29	83,723	95086,428	5299153,836	55,705
22	0,00091	0,00091	95044,567	86,199	95001,467	5204067,408	54,754
23	0,00094	0,00094	94958,368	89,334	94913,701	5109065,941	53,803
24	0,00099	0,00099	94869,034	93,585	94822,241	5014152,24	52,853
25	0,00098	0,00098	94775,448	92,554	94729,171	4919329,999	51,905
26	0,00103	0,00103	94682,894	97,15	94634,319	4824600,828	50,955
27	0,00108	0,00108	94585,744	102,4	94534,543	4729966,509	50,007
28	0,0011	0,0011	94483,343	103,617	94431,535	4635431,965	49,061
29	0,00112	0,00112	94379,726	105,776	94326,838	4541000,431	48,114
30	0,00122	0,00122	94273,95	115,207	94216,346	4446673,593	47,168
31	0,00121	0,00121	94158,743	113,867	94101,809	4352457,247	46,225
32	0,0013	0,0013	94044,876	122,268	93983,742	4258355,438	45,28
33	0,0014	0,0014	93922,608	131,475	93856,87	4164371,696	44,338
34	0,00145	0,00144	93791,133	135,491	93723,387	4070514,826	43,4
35	0,00157	0,00157	93655,641	147,07	93582,106	3976791,439	42,462
36	0,00163	0,00163	93508,571	152,722	93432,21	3883209,332	41,528
37	0,00171	0,00171	93355,849	159,429	93276,135	3789777,122	40,595
38	0,0019	0,0019	93196,42	176,625	93108,107	3696500,987	39,664

Continuação da Tabela D.4

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
39	0,00197	0,00197	93019,795	182,974	92928,308	3603392,88	38,738
40	0,00219	0,00219	92836,821	202,987	92735,328	3510464,572	37,813
41	0,0021	0,0021	92633,834	194,123	92536,773	3417729,244	36,895
42	0,00246	0,00246	92439,712	226,942	92326,241	3325192,471	35,971
43	0,00247	0,00247	92212,77	227,606	92098,967	3232866,231	35,059
44	0,00265	0,00264	91985,164	243,162	91863,583	3140767,264	34,144
45	0,00294	0,00293	91742,002	269,083	91607,46	3048903,681	33,233
46	0,00305	0,00305	91472,919	278,571	91333,633	2957296,221	32,33
47	0,00322	0,00322	91194,348	293,196	91047,75	2865962,588	31,427
48	0,00368	0,00367	90901,152	334,054	90734,125	2774914,838	30,527
49	0,00371	0,00371	90567,098	335,679	90399,258	2684180,713	29,637
50	0,00431	0,0043	90231,419	388,159	90037,339	2593781,455	28,746
51	0,0042	0,00419	89843,259	376,336	89655,091	2503744,116	27,868
52	0,00491	0,0049	89466,923	438,626	89247,61	2414089,025	26,983
53	0,00514	0,00512	89028,297	456,244	88800,174	2324841,416	26,114
54	0,00544	0,00543	88572,052	480,591	88331,756	2236041,241	25,245
55	0,00601	0,00599	88091,461	527,571	87827,675	2147709,485	24,38
56	0,00651	0,00648	87563,889	567,831	87279,974	2059881,81	23,524
57	0,00679	0,00677	86996,059	589,045	86701,536	1972601,836	22,675
58	0,00773	0,0077	86407,014	665,445	86074,291	1885900,3	21,826
59	0,00792	0,00789	85741,569	676,538	85403,3	1799826,008	20,991
60	0,00997	0,00992	85065,031	843,693	84643,185	1714422,708	20,154
61	0,00924	0,0092	84221,338	774,93	83833,873	1629779,524	19,351
62	0,01117	0,0111	83446,407	926,521	82983,147	1545945,651	18,526
63	0,01225	0,01218	82519,886	1004,934	82017,419	1462962,504	17,729
64	0,01358	0,01349	81514,952	1099,481	80965,211	1380945,085	16,941
65	0,01556	0,01544	80415,471	1241,664	79794,639	1299979,874	16,166
66	0,01678	0,01664	79173,807	1317,617	78514,998	1220185,235	15,411
67	0,0184	0,01824	77856,19	1419,858	77146,261	1141670,237	14,664

Continuação da Tabela D.4

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
68	0,02144	0,02122	76436,332	1621,759	75625,453	1064523,976	13,927
69	0,022	0,02176	74814,573	1628,148	74000,499	988898,523	13,218
70	0,0292	0,02878	73186,425	2106,543	72133,154	914898,024	12,501
71	0,02743	0,02706	71079,882	1923,358	70118,203	842764,87	11,857
72	0,03464	0,03405	69156,524	2354,802	67979,123	772646,667	11,172
73	0,03802	0,03731	66801,722	2492,515	65555,464	704667,544	10,549
74	0,04356	0,04263	64309,206	2741,726	62938,343	639112,08	9,938
75	0,05168	0,05038	61567,48	3101,532	60016,714	576173,736	9,358
76	0,05563	0,05412	58465,948	3164,2	56883,848	516157,022	8,828
77	0,06206	0,06019	55301,748	3328,819	53637,339	459273,174	8,305
78	0,07392	0,07128	51972,929	3704,86	50120,499	405635,835	7,805
79	0,07552	0,07277	48268,07	3512,348	46511,895	355515,336	7,365
80	0,09146	0,08746	44755,721	3914,374	42798,534	309003,44	6,904
81	0,08494	0,08148	40841,347	3327,63	39177,532	266204,906	6,518
82	0,10436	0,09918	37513,718	3720,701	35653,367	227027,374	6,052
83	0,1123	0,10633	33793,016	3593,131	31996,45	191374,007	5,663
84	0,12857	0,12081	30199,885	3648,342	28375,714	159377,556	5,277
85	0,14922	0,13886	26551,543	3686,937	24708,075	131001,842	4,934
86	0,15867	0,14701	22864,606	3361,356	21183,928	106293,768	4,649
87	0,1714	0,15787	19503,25	3078,984	17963,758	85109,84	4,364
88	0,19521	0,17785	16424,266	2921,03	14963,751	67146,082	4,088
89	0,19743	0,17969	13503,236	2426,397	12290,038	52182,33	3,864
90	0,24564	0,21877	11076,84	2423,263	9865,208	39892,292	3,601
91	0,21564	0,19465	8653,577	1684,447	7811,353	30027,084	3,47
92	0,25415	0,2255	6969,13	1571,508	6183,375	22215,731	3,188
93	0,26466	0,23373	5397,621	1261,589	4766,827	16032,355	2,97
94	0,28486	0,24935	4136,033	1031,303	3620,381	11265,528	2,724
95	0,33494	0,28689	3104,729	890,729	2659,365	7645,147	2,462
96	0,35141	0,29889	2214,001	661,75	1883,126	4985,782	2,252

Continuação da Tabela D.4

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
97	0,37402	0,31509	1552,251	489,1	1307,701	3102,657	1,999
98	0,44216	0,3621	1063,151	384,97	870,666	1794,956	1,688
99	0,41381	0,34287	678,181	232,526	561,918	924,29	1,363
100	0,45965	0,37375	445,655	166,565	362,372	0	0

Fim da Tabela

Tabela D.5: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sem Distinção de Sexo.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,05194	0,05062	0,00001	5062,494	97468,753	6667713,816	66,677
1	0,00886	0,00882	94937,506	837,331	94518,841	6570245,063	69,206
2	0,00404	0,00403	94100,176	379,019	93910,666	6475726,222	68,817
3	0,00236	0,00236	93721,157	221,304	93610,504	6381815,556	68,094
4	0,00172	0,00172	93499,852	161,077	93419,314	6288205,052	67,254
5	0,00131	0,00131	93338,776	121,841	93277,855	6194785,738	66,369
6	0,0011	0,0011	93216,935	102,686	93165,592	6101507,882	65,455
7	0,001	0,001	93114,249	92,831	93067,833	6008342,291	64,527
8	0,00085	0,00085	93021,418	78,954	92981,94	5915274,457	63,59
9	0,00079	0,00078	92942,463	72,954	92905,986	5822292,517	62,644
10	0,00074	0,00074	92869,509	68,658	92835,18	5729386,531	61,693
11	0,00072	0,00072	92800,851	67,225	92767,238	5636551,351	60,738
12	0,00072	0,00072	92733,626	66,376	92700,438	5543784,113	59,782
13	0,00076	0,00076	92667,25	70,826	92631,837	5451083,675	58,824
14	0,00088	0,00088	92596,424	81,367	92555,74	5358451,838	57,869
15	0,00101	0,00101	92515,056	93,781	92468,166	5265896,098	56,919
16	0,00122	0,00122	92421,275	112,988	92364,781	5173427,933	55,977

Continuação da Tabela D.5

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
17	0,00138	0,00138	92308,287	127,402	92244,586	5081063,151	55,044
18	0,0016	0,0016	92180,884	147,468	92107,15	4988818,566	54,12
19	0,00172	0,00171	92033,416	157,756	91954,538	4896711,416	53,206
20	0,0018	0,0018	91875,66	165,634	91792,843	4804756,877	52,296
21	0,00185	0,00185	91710,026	169,836	91625,108	4712964,034	51,39
22	0,00191	0,00191	91540,19	174,97	91452,705	4621338,926	50,484
23	0,00194	0,00193	91365,22	176,708	91276,866	4529886,221	49,58
24	0,002	0,00199	91188,512	181,779	91097,622	4438609,355	48,675
25	0,00198	0,00198	91006,733	180,129	90916,668	4347511,733	47,771
26	0,00203	0,00203	90826,603	183,945	90734,631	4256595,065	46,865
27	0,00208	0,00207	90642,659	187,963	90548,677	4165860,435	45,959
28	0,00215	0,00215	90454,696	194,608	90357,392	4075311,757	45,054
29	0,0021	0,0021	90260,088	189,403	90165,386	3984954,366	44,15
30	0,00226	0,00226	90070,685	203,557	89968,906	3894788,979	43,241
31	0,00219	0,00218	89867,128	196,162	89769,047	3804820,073	42,338
32	0,00239	0,00239	89670,966	213,996	89563,968	3715051,026	41,43
33	0,00252	0,00252	89456,97	225,22	89344,36	3625487,058	40,528
34	0,00257	0,00256	89231,75	228,862	89117,319	3536142,699	39,629
35	0,00275	0,00274	89002,888	244,308	88880,734	3447025,38	38,729
36	0,00286	0,00285	88758,58	253,342	88631,909	3358144,646	37,835
37	0,00297	0,00296	88505,238	262,182	88374,146	3269512,737	36,941
38	0,00324	0,00324	88243,055	285,502	88100,304	3181138,591	36,05
39	0,00326	0,00326	87957,553	286,691	87814,207	3093038,287	35,165
40	0,00366	0,00365	87670,861	319,959	87510,882	3005224,08	34,278
41	0,00347	0,00346	87350,902	302,201	87199,801	2917713,198	33,402
42	0,00411	0,0041	87048,701	356,847	86870,277	2830513,397	32,516
43	0,00411	0,0041	86691,854	355,796	86513,956	2743643,119	31,648
44	0,00444	0,00443	86336,057	382,597	86144,759	2657129,164	30,777
45	0,00489	0,00488	85953,46	419,354	85743,783	2570984,405	29,911

Continuação da Tabela D.5

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
46	0,00496	0,00494	85534,106	422,839	85322,687	2485240,622	29,056
47	0,00525	0,00524	85111,267	445,682	84888,426	2399917,935	28,197
48	0,00598	0,00596	84665,585	504,646	84413,262	2315029,509	27,343
49	0,00611	0,00609	84160,939	512,393	83904,742	2230616,247	26,504
50	0,00702	0,00699	83648,546	584,939	83356,076	2146711,505	25,663
51	0,00669	0,00667	83063,607	554,236	82786,489	2063355,429	24,841
52	0,00789	0,00785	82509,371	648,081	82185,331	1980568,94	24,004
53	0,00834	0,0083	81861,29	679,5	81521,54	1898383,609	23,19
54	0,00891	0,00887	81181,791	720,447	80821,567	1816862,069	22,38
55	0,00976	0,00972	80461,343	781,882	80070,402	1736040,502	21,576
56	0,01054	0,01049	79679,461	835,665	79261,628	1655970,099	20,783
57	0,01106	0,011	78843,796	867,419	78410,086	1576708,471	19,998
58	0,01244	0,01236	77976,377	964,006	77494,374	1498298,385	19,215
59	0,01289	0,0128	77012,371	986,088	76519,327	1420804,011	18,449
60	0,01579	0,01567	76026,283	1191,258	75430,654	1344284,684	17,682
61	0,01477	0,01466	74835,025	1097,444	74286,303	1268854,031	16,955
62	0,01769	0,01754	73737,581	1293,309	73090,927	1194567,728	16,2
63	0,01936	0,01918	72444,272	1389,197	71749,673	1121476,801	15,481
64	0,02113	0,02091	71055,075	1485,481	70312,334	1049727,128	14,773
65	0,02385	0,02357	69569,594	1639,643	68749,773	979414,793	14,078
66	0,0257	0,02537	67929,951	1723,543	67068,18	910665,021	13,406
67	0,02786	0,02747	66206,408	1818,861	65296,978	843596,841	12,742
68	0,03205	0,03155	64387,547	2031,258	63371,918	778299,863	12,088
69	0,0329	0,03237	62356,29	2018,408	61347,086	714927,945	11,465
70	0,04222	0,04135	60337,882	2494,788	59090,488	653580,859	10,832
71	0,03988	0,0391	57843,094	2261,435	56712,377	594490,371	10,278
72	0,04927	0,04809	55581,66	2672,782	54245,269	537777,994	9,675
73	0,0535	0,0521	52908,878	2756,777	51530,489	483532,725	9,139
74	0,06027	0,0585	50152,1	2934,088	48685,057	432002,236	8,614

Continuação da Tabela D.5

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
75	0,06991	0,06755	47218,013	3189,381	45623,322	383317,179	8,118
76	0,07477	0,07208	44028,632	3173,564	42441,85	337693,857	7,67
77	0,08313	0,07981	40855,068	3260,598	39224,769	295252,007	7,227
78	0,0969	0,09242	37594,47	3474,625	35857,157	256027,239	6,81
79	0,09892	0,09426	34119,845	3216,008	32511,841	220170,082	6,453
80	0,11475	0,10852	30903,837	3353,657	29227,009	187658,241	6,072
81	0,10872	0,10312	27550,18	2840,875	26129,742	158431,232	5,751
82	0,12937	0,12151	24709,305	3002,429	23208,09	132301,49	5,354
83	0,13892	0,12989	21706,875	2819,612	20297,069	109093,4	5,026
84	0,15684	0,14543	18887,263	2746,858	17513,834	88796,33	4,701
85	0,17688	0,16251	16140,405	2622,998	14828,906	71282,496	4,416
86	0,18756	0,17148	13517,407	2317,987	12358,413	56453,59	4,176
87	0,20149	0,18305	11199,419	2050,006	10174,416	44095,177	3,937
88	0,2266	0,20354	9149,414	1862,272	8218,277	33920,761	3,707
89	0,22992	0,20621	7287,141	1502,7	6535,791	25702,483	3,527
90	0,27222	0,2396	5784,441	1385,979	5091,452	19166,692	3,313
91	0,24932	0,22168	4398,462	975,059	3910,933	14075,24	3,2
92	0,28615	0,25033	3423,403	856,984	2994,911	10164,308	2,969
93	0,29579	0,25768	2566,42	661,314	2235,762	7169,396	2,794
94	0,31181	0,26975	1905,105	513,908	1648,151	4933,634	2,59
95	0,35585	0,3021	1391,197	420,277	1181,058	3285,483	2,362
96	0,37649	0,31684	970,92	307,629	817,105	2104,424	2,167
97	0,39338	0,32872	663,291	218,037	554,272	1287,319	1,941
98	0,46716	0,3787	445,254	168,619	360,944	733,047	1,646
99	0,43156	0,35496	276,635	98,195	227,538	372,102	1,345
100	0,46866	0,37969	178,44	67,751	144,565	0	0

Fim da Tabela

Tabela D.6: Tábua de Mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sem Distinção de Sexo.

Início da Tabela							
Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
0	0,03472	0,03413	0,00001	3413,039	98293,48	6957026,355	69,57
1	0,00559	0,00557	96586,961	538,435	96317,744	6858732,875	71,011
2	0,00278	0,00278	96048,526	266,554	95915,25	6762415,131	70,406
3	0,0017	0,0017	95781,973	162,857	95700,544	6666499,882	69,601
4	0,00126	0,00126	95619,116	120,343	95558,944	6570799,337	68,718
5	0,00098	0,00098	95498,773	93,253	95452,146	6475240,393	67,804
6	0,00083	0,00083	95405,52	78,757	95366,141	6379788,247	66,87
7	0,00076	0,00076	95326,763	72,35	95290,588	6284422,106	65,925
8	0,00065	0,00065	95254,413	62,234	95223,296	6189131,518	64,975
9	0,00061	0,00061	95192,179	57,6	95163,379	6093908,222	64,017
10	0,00058	0,00058	95134,579	55,1	95107,028	5998744,843	63,055
11	0,00057	0,00057	95079,478	54,255	95052,351	5903637,815	62,092
12	0,00056	0,00056	95025,223	53,411	94998,518	5808585,464	61,127
13	0,00059	0,00059	94971,812	56,49	94943,567	5713586,947	60,161
14	0,00069	0,00069	94915,322	65,711	94882,467	5618643,38	59,196
15	0,00081	0,00081	94849,611	77,106	94811,058	5523760,913	58,237
16	0,00098	0,00098	94772,504	92,42	94726,294	5428949,855	57,284
17	0,00109	0,00109	94680,084	103,39	94628,389	5334223,561	56,339
18	0,00127	0,00127	94576,694	120,124	94516,632	5239595,172	55,4
19	0,00137	0,00137	94456,57	129,434	94391,853	5145078,54	54,47
20	0,00144	0,00144	94327,136	135,772	94259,25	5050686,687	53,544
21	0,00148	0,00148	94191,364	139,168	94121,78	4956427,438	52,621
22	0,0015	0,0015	94052,196	140,857	93981,767	4862305,658	51,698
23	0,00153	0,00153	93911,338	143,501	93839,588	4768323,891	50,775
24	0,00159	0,00159	93767,838	148,97	93693,353	4674484,303	49,852
25	0,00159	0,00159	93618,868	148,789	93544,473	4580790,95	48,93

Continuação da Tabela D.6

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
26	0,00164	0,00163	93470,079	152,776	93393,691	4487246,477	48,007
27	0,00168	0,00168	93317,303	156,773	93238,917	4393852,785	47,085
28	0,00175	0,00174	93160,53	162,531	93079,265	4300613,869	46,163
29	0,00174	0,00174	92998	161,412	92917,293	4207534,604	45,243
30	0,00186	0,00186	92836,587	172,442	92750,366	4114617,31	44,321
31	0,00184	0,00184	92664,145	170,3	92578,996	4021866,944	43,403
32	0,00198	0,00198	92493,846	183,287	92402,202	3929287,948	42,482
33	0,0021	0,0021	92310,559	193,413	92213,852	3836885,746	41,565
34	0,00217	0,00217	92117,145	199,694	92017,298	3744671,894	40,651
35	0,00232	0,00232	91917,452	213,333	91810,785	3652654,596	39,738
36	0,00242	0,00242	91704,118	221,565	91593,335	3560843,811	38,83
37	0,00254	0,00253	91482,553	231,793	91366,656	3469250,475	37,923
38	0,00277	0,00276	91250,76	252,161	91124,679	3377883,819	37,018
39	0,00286	0,00286	90998,598	259,887	90868,655	3286759,14	36,119
40	0,00317	0,00316	90738,711	287,172	90595,125	3195890,485	35,221
41	0,00311	0,00311	90451,539	281,241	90310,919	3105295,36	34,331
42	0,00361	0,0036	90170,298	324,963	90007,817	3014984,441	33,437
43	0,00369	0,00369	89845,335	331,12	89679,775	2924976,624	32,556
44	0,00398	0,00398	89514,215	355,888	89336,271	2835296,849	31,674
45	0,00438	0,00437	89158,327	389,295	88963,68	2745960,578	30,799
46	0,00452	0,00451	88769,032	400,283	88568,891	2656996,898	29,932
47	0,00482	0,00481	88368,749	424,731	88156,384	2568428,008	29,065
48	0,00544	0,00543	87944,018	477,378	87705,329	2480271,624	28,203
49	0,00562	0,00561	87466,64	490,458	87221,411	2392566,294	27,354
50	0,00639	0,00637	86976,182	553,864	86699,251	2305344,883	26,505
51	0,00624	0,00622	86422,319	537,485	86153,576	2218645,632	25,672
52	0,00724	0,00722	85884,834	619,75	85574,959	2132492,056	24,83
53	0,0077	0,00767	85265,084	654,261	84937,954	2046917,097	24,007
54	0,0082	0,00817	84610,823	691,101	84265,273	1961979,143	23,188

Continuação da Tabela D.6

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
55	0,00898	0,00894	83919,722	750,645	83544,4	1877713,87	22,375
56	0,00968	0,00963	83169,078	800,888	82768,634	1794169,47	21,573
57	0,01021	0,01016	82368,19	836,948	81949,716	1711400,836	20,777
58	0,01141	0,01135	81531,242	925,361	81068,562	1629451,12	19,986
59	0,01187	0,0118	80605,882	951,362	80130,201	1548382,558	19,209
60	0,0143	0,0142	79654,519	1131,048	79088,995	1468252,358	18,433
61	0,01363	0,01354	78523,471	1062,889	77992,027	1389163,362	17,691
62	0,01615	0,01602	77460,582	1240,757	76840,204	1311171,336	16,927
63	0,01763	0,01748	76219,825	1331,962	75553,844	1234331,132	16,194
64	0,01922	0,01903	74887,863	1425,365	74175,181	1158777,288	15,473
65	0,02162	0,02139	73462,498	1571,303	72676,847	1084602,107	14,764
66	0,02329	0,02302	71891,195	1654,937	71063,727	1011925,26	14,076
67	0,02527	0,02495	70236,258	1752,454	69360,031	940861,534	13,396
68	0,02893	0,02852	68483,804	1953,058	67507,275	871501,503	12,726
69	0,02973	0,0293	66530,746	1949,093	65556,2	803994,227	12,085
70	0,03772	0,03702	64581,653	2391,02	63386,143	738438,028	11,434
71	0,03599	0,03535	62190,633	2198,506	61091,38	675051,885	10,855
72	0,04433	0,04337	59992,128	2601,571	58691,342	613960,504	10,234
73	0,04817	0,04704	57390,557	2699,432	56040,841	555269,162	9,675
74	0,05415	0,05273	54691,125	2883,604	53249,323	499228,321	9,128
75	0,06281	0,06089	51807,522	3154,761	50230,141	445978,998	8,608
76	0,06719	0,065	48652,761	3162,593	47071,465	395748,856	8,134
77	0,07481	0,07211	45490,168	3280,405	43849,966	348677,391	7,665
78	0,08717	0,08353	42209,763	3525,934	40446,796	304827,426	7,222
79	0,08867	0,0849	38683,829	3284,394	37041,632	264380,63	6,834
80	0,10442	0,09924	35399,435	3513,146	33642,862	227338,998	6,422
81	0,0984	0,09379	31886,289	2990,596	30390,991	193696,136	6,075
82	0,11862	0,11197	28895,693	3235,569	27277,908	163305,146	5,652
83	0,12733	0,11971	25660,123	3071,739	24124,254	136027,238	5,301

Continuação da Tabela D.6

Idade	$\dot{m}_{x,1970}$	$\dot{q}_{x,1970}$	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{L}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$
84	0,14448	0,13475	22588,385	3043,687	21066,541	111902,984	4,954
85	0,16436	0,15188	19544,698	2968,487	18060,454	90836,443	4,648
86	0,17418	0,16022	16576,211	2655,907	15248,257	72775,988	4,39
87	0,18734	0,1713	13920,304	2384,513	12728,047	57527,731	4,133
88	0,2125	0,19209	11535,791	2215,939	10427,822	44799,684	3,884
89	0,2139	0,19324	9319,852	1800,938	8419,383	34371,862	3,688
90	0,26006	0,23013	7518,914	1730,356	6653,736	25952,479	3,452
91	0,23018	0,20642	5788,558	1194,868	5191,125	19298,743	3,334
92	0,2705	0,23828	4593,691	1094,569	4046,406	14107,618	3,071
93	0,28053	0,24602	3499,122	860,868	3068,688	10061,212	2,875
94	0,29878	0,25995	2638,254	685,811	2295,349	6992,524	2,65
95	0,34545	0,29457	1952,443	575,127	1664,88	4697,175	2,406
96	0,36573	0,30919	1377,317	425,854	1164,389	3032,295	2,202
97	0,38489	0,32277	951,462	307,104	797,91	1867,906	1,963
98	0,45685	0,3719	644,358	239,634	524,541	1069,996	1,661
99	0,42747	0,35219	404,724	142,54	333,454	545,455	1,348
100	0,47341	0,3828	262,183	100,363	212,002	0	0

Fim da Tabela

Apêndice E

Cálculo das Taxas Vitalícias

Tabela E.1: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Masculino.

Início da Tabela					
Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
0	100000	5450,348	6342711,384	63,427	63,969
1	94549,652	871,569	6247344,179	66,075	66,617
2	93678,083	406,721	6153230,312	65,685	66,227
3	93271,362	239,749	6059755,589	64,969	65,511
4	93031,613	175,879	5966604,102	64,135	64,677
5	92855,734	134,684	5873660,428	63,256	63,798
6	92721,051	113,192	5780872,036	62,347	62,889
7	92607,858	106,128	5688207,581	61,423	61,965
8	92501,731	89,763	5595652,787	60,492	61,034
9	92411,968	83,683	5503195,937	59,551	60,093
10	92328,284	78,915	5410825,812	58,604	59,146
11	92249,37	77,581	5318536,985	57,654	58,196
12	92171,789	75,502	5226326,405	56,702	57,244
13	92096,286	80,987	5134192,368	55,748	56,29
14	92015,299	94,101	5042136,575	54,797	55,339
15	91921,198	109,813	4950168,326	53,852	54,394

Continuação da Tabela E.1

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
16	91811,384	137,688	4858302,035	52,916	53,458
17	91673,696	160,591	4766559,495	51,995	52,537
18	91513,105	188,482	4674966,095	51,085	51,627
19	91324,622	204,711	4583547,231	50,19	50,732
20	91119,911	214,065	4492324,964	49,301	49,843
21	90905,846	221,22	4401312,086	48,416	48,958
22	90684,625	225,338	4310516,851	47,533	48,075
23	90459,288	227,941	4219944,894	46,65	47,192
24	90231,347	233,474	4129599,577	45,767	46,309
25	89997,873	230,88	4039484,967	44,884	45,426
26	89766,993	232,967	3949602,534	43,998	44,54
27	89534,026	236,348	3859952,024	43,112	43,654
28	89297,678	248,072	3770536,172	42,224	42,766
29	89049,606	241,07	3681362,53	41,341	41,883
30	88808,536	255,495	3592433,459	40,451	40,993
31	88553,041	250,963	3503752,671	39,567	40,109
32	88302,078	272,715	3415325,112	38,678	39,22
33	88029,363	285,75	3327159,391	37,796	38,338
34	87743,613	290,912	3239272,903	36,917	37,459
35	87452,701	310,481	3151674,746	36,039	36,581
36	87142,22	322,514	3064377,285	35,165	35,707
37	86819,705	335,476	2977396,323	34,294	34,836
38	86484,23	362,05	2890744,355	33,425	33,967
39	86122,18	364,863	2804441,15	32,564	33,106
40	85757,317	405,69	2718501,402	31,7	32,242
41	85351,627	391,618	2632946,93	30,848	31,39
42	84960,009	458,652	2547791,111	29,988	30,53
43	84501,358	461,968	2463060,428	29,148	29,69
44	84039,389	500,464	2378790,055	28,306	28,848

Continuação da Tabela E.1

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
45	83538,926	545,004	2295000,897	27,472	28,014
46	82993,921	547,343	2211734,474	26,649	27,191
47	82446,578	582,067	2129014,224	25,823	26,365
48	81864,512	651,878	2046858,679	25,003	25,545
49	81212,634	671,972	1965320,106	24,2	24,742
50	80540,662	753,457	1884443,458	23,397	23,939
51	79787,205	717,289	1804279,524	22,614	23,156
52	79069,916	829,43	1724850,963	21,814	22,356
53	78240,486	878,708	1646195,762	21,04	21,582
54	77361,778	934,579	1568394,63	20,274	20,816
55	76427,199	1006,873	1491500,142	19,515	20,057
56	75420,325	1070,949	1415576,38	18,769	19,311
57	74349,377	1118,159	1340691,529	18,032	18,574
58	73231,218	1223,353	1266901,231	17,3	17,842
59	72007,865	1267,188	1194281,69	16,585	17,127
60	70740,677	1464,033	1122907,419	15,874	16,416
61	69276,644	1388,991	1052898,759	15,198	15,74
62	67887,653	1597,892	984316,61	14,499	15,041
63	66289,761	1707,459	917227,904	13,837	14,379
64	64582,302	1800,902	851791,872	13,189	13,731
65	62781,401	1943,063	788110,021	12,553	13,095
66	60838,337	2039,397	726300,152	11,938	12,48
67	58798,94	2125,584	666481,513	11,335	11,877
68	56673,356	2324,561	608745,365	10,741	11,283
69	54348,795	2312,326	553234,289	10,179	10,721
70	52036,47	2694,629	500041,657	9,609	10,151
71	49341,841	2496,541	449352,502	9,107	9,649
72	46845,3	2847,531	401258,931	8,566	9,108
73	43997,769	2885,547	355837,396	8,088	8,63

Continuação da Tabela E.1

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
74	41112,222	2987,701	313282,401	7,62	8,162
75	38124,521	3123,886	273664,029	7,178	7,72
76	35000,635	3073,52	237101,451	6,774	7,316
77	31927,114	3115,388	203637,577	6,378	6,92
78	28811,726	3184,845	173268,157	6,014	6,556
79	25626,881	2909,527	146048,853	5,699	6,241
80	22717,354	2880,647	121876,735	5,365	5,907
81	19836,707	2480,416	100599,705	5,071	5,613
82	17356,291	2499,969	82003,206	4,725	5,267
83	14856,322	2297,433	65896,9	4,436	4,978
84	12558,889	2155,095	52189,294	4,156	4,698
85	10403,794	1950,627	40707,952	3,913	4,455
86	8453,167	1667,34	31279,471	3,7	4,242
87	6785,827	1441,147	23659,974	3,487	4,029
88	5344,68	1254,7	17594,721	3,292	3,834
89	4089,98	969,279	12877,391	3,149	3,691
90	3120,701	826,862	9272,051	2,971	3,513
91	2293,839	583,018	6564,781	2,862	3,404
92	1710,821	488,96	4562,451	2,667	3,209
93	1221,861	360,972	3096,109	2,534	3,076
94	860,889	258,94	2054,734	2,387	2,929
95	601,949	195,182	1323,315	2,198	2,74
96	406,767	141,831	818,957	2,013	2,555
97	264,936	93,985	483,105	1,823	2,365
98	170,951	70,95	265,162	1,551	2,093
99	100,001	38,626	129,686	1,297	1,839
100	61,375	24,756	0	0	0,542

Fim da Tabela

Tabela E.2: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Masculino.

Início da Tabela					
Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
0	100000	3832,888	6602040,677	66,02	66,562
1	96167,112	594,738	6505298,631	67,646	68,188
2	95572,373	299,653	6409428,889	67,064	67,606
3	95272,72	182,594	6314006,342	66,273	66,815
4	95090,127	135,85	6218824,918	65,399	65,941
5	94954,276	105,833	6123802,717	64,492	65,034
6	94848,443	88,182	6028901,357	63,564	64,106
7	94760,262	84,592	5934097,005	62,622	63,164
8	94675,67	71,732	5839379,039	61,678	62,22
9	94603,937	67,439	5744739,236	60,724	61,266
10	94536,498	64,224	5650169,018	59,767	60,309
11	94472,274	64,035	5555664,632	58,807	59,349
12	94408,239	62,433	5461224,376	57,847	58,389
13	94345,806	67,087	5366847,353	56,885	57,427
14	94278,72	79,215	5272535,09	55,925	56,467
15	94199,505	94,485	5178295,978	54,972	55,514
16	94105,02	118,381	5084143,716	54,026	54,568
17	93986,638	137,476	4990097,887	53,094	53,636
18	93849,162	162,346	4896179,987	52,171	52,713
19	93686,816	176,923	4802411,997	51,26	51,802
20	93509,893	184,814	4708813,643	50,356	50,898
21	93325,08	189,801	4615396,156	49,455	49,997
22	93135,278	190,72	4522165,977	48,555	49,097
23	92944,558	193,542	4429126,058	47,653	48,195
24	92751,017	201,363	4336278,271	46,752	47,294

Continuação da Tabela E.2

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
25	92549,653	200,913	4243627,936	45,852	46,394
26	92348,741	204,194	4151178,739	44,951	45,493
27	92144,547	207,366	4058932,095	44,05	44,592
28	91937,181	218,028	3966891,231	43,148	43,69
29	91719,153	214,48	3875063,064	42,249	42,791
30	91504,673	227,125	3783451,151	41,347	41,889
31	91277,548	226,081	3692060,041	40,449	40,991
32	91051,467	243,479	3600895,533	39,548	40,09
33	90807,988	255,431	3509965,806	38,653	39,195
34	90552,557	263,828	3419285,534	37,76	38,302
35	90288,729	280,558	3328864,891	36,869	37,411
36	90008,171	291,725	3238716,441	35,982	36,524
37	89716,446	306,062	3148854,132	35,098	35,64
38	89410,383	329,848	3059290,718	34,216	34,758
39	89080,535	339,553	2970045,259	33,341	33,883
40	88740,982	374,603	2881134,5	32,467	33,009
41	88366,38	372,294	2792580,819	31,602	32,144
42	87994,086	428,129	2704400,586	30,734	31,276
43	87565,956	439,437	2616620,565	29,882	30,424
44	87126,52	474,72	2529274,327	29,03	29,572
45	86651,799	516,445	2442385,168	28,186	28,728
46	86135,355	528,034	2355991,591	27,352	27,894
47	85607,321	564,205	2270120,253	26,518	27,06
48	85043,116	628,745	2184795,034	25,69	26,232
49	84414,37	654,439	2100066,291	24,878	25,42
50	83759,931	728,939	2015979,14	24,069	24,611
51	83030,993	706,838	1932583,678	23,275	23,817
52	82324,154	810,088	1849906,105	22,471	23,013
53	81514,066	862,649	1767986,995	21,689	22,231

Continuação da Tabela E.2

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
54	80651,417	913,103	1686904,254	20,916	21,458
55	79738,313	985,628	1606709,389	20,15	20,692
56	78752,686	1046,741	1527463,889	19,396	19,938
57	77705,944	1097,589	1449234,574	18,65	19,192
58	76608,355	1198,136	1372077,425	17,91	18,452
59	75410,219	1241,443	1296068,138	17,187	17,729
60	74168,776	1426,24	1221278,641	16,466	17,008
61	72742,535	1364,193	1147822,985	15,779	16,321
62	71378,342	1564,793	1075762,547	15,071	15,613
63	69813,549	1670,297	1005166,601	14,398	14,94
64	68143,252	1762,951	936188,2	13,739	14,281
65	66380,3	1904,186	868926,424	13,09	13,632
66	64476,114	1999,869	803498,217	12,462	13,004
67	62476,245	2089,309	740022,038	11,845	12,387
68	60386,936	2285,879	678590,447	11,237	11,779
69	58101,057	2275,323	619346,45	10,66	11,202
70	55825,734	2657,555	562383,055	10,074	10,616
71	53168,179	2472,451	507886,098	9,552	10,094
72	50695,728	2839,587	455954,145	8,994	9,536
73	47856,141	2892,694	406678,211	8,498	9,04
74	44963,447	3008,209	360268,417	8,012	8,554
75	41955,237	3175,544	316809,075	7,551	8,093
76	38779,693	3133,818	276441,61	7,129	7,671
77	35645,875	3208,306	239228,826	6,711	7,253
78	32437,569	3313,465	205187,104	6,326	6,868
79	29124,104	3032,148	174406,267	5,988	6,53
80	26091,956	3089,784	146798,237	5,626	6,168
81	23002,172	2654,075	122251,173	5,315	5,857
82	20348,096	2746,057	100576,039	4,943	5,485

Continuação da Tabela E.2					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
83	17602,039	2549,09	81600,971	4,636	5,178
84	15052,949	2437,901	65273,477	4,336	4,878
85	12615,048	2245,984	51439,478	4,078	4,62
86	10369,064	1948,248	39947,422	3,853	4,395
87	8420,816	1694,46	30552,482	3,628	4,17
88	6726,356	1515,327	22978,896	3,416	3,958
89	5211,029	1181,554	17010,204	3,264	3,806
90	4029,475	1040,389	12389,952	3,075	3,617
91	2989,086	715,294	8880,671	2,971	3,513
92	2273,792	625,604	6249,232	2,748	3,29
93	1648,189	468,261	4288,242	2,602	3,144
94	1179,928	346,726	2874,183	2,436	2,978
95	833,202	265,415	1867,619	2,241	2,783
96	567,787	194,585	1167,124	2,056	2,598
97	373,202	129,144	696,629	1,867	2,409
98	244,059	97,004	387,999	1,59	2,132
99	147,054	55,982	192,443	1,309	1,851
100	91,072	35,386	0	0	0,542
Fim da Tabela					

Tabela E.3: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sexo Feminino.

Início da Tabela					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
0	100000	4491,955	7000690,988	70,007	70,549
1	95508,045	806,695	6904464,23	72,292	72,834
2	94701,35	352,51	6809359,533	71,904	72,446

Continuação da Tabela E.3

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
3	94348,84	203,724	6714834,437	71,17	71,712
4	94145,116	146,191	6620587,459	70,323	70,865
5	93998,925	108,814	6526515,439	69,432	69,974
6	93890,111	92,357	6432570,921	68,512	69,054
7	93797,754	78,972	6338726,989	67,579	68,121
8	93718,782	67,914	6244968,721	66,635	67,177
9	93650,867	61,525	6151283,896	65,683	66,225
10	93589,343	57,947	6057663,791	64,726	65,268
11	93531,396	56,437	5964103,422	63,766	64,308
12	93474,959	56,253	5870600,245	62,804	63,346
13	93418,706	59,731	5777153,412	61,842	62,384
14	93358,976	67,295	5683764,571	60,881	61,423
15	93291,681	75,027	5590439,242	59,924	60,466
16	93216,654	83,58	5497185,074	58,972	59,514
17	93133,075	87,369	5404010,21	58,025	58,567
18	93045,706	97,121	5310920,82	57,079	57,621
19	92948,584	102,432	5217923,675	56,138	56,68
20	92846,153	108,876	5125026,307	55,199	55,741
21	92737,276	111,07	5032234,592	54,263	54,805
22	92626,206	115,594	4939552,851	53,328	53,87
23	92510,612	118,353	4846984,442	52,394	52,936
24	92392,259	124,148	4754533,006	51,46	52,002
25	92268,111	122,356	4662202,822	50,529	51,071
26	92145,755	127,853	4569995,889	49,595	50,137
27	92017,901	133,494	4477914,061	48,664	49,206
28	91884,408	135,218	4385962,906	47,733	48,275
29	91749,19	134,164	4294146,108	46,803	47,345
30	91615,026	147,377	4202464	45,871	46,413
31	91467,649	140,229	4110922,662	44,944	45,486

Continuação da Tabela E.3

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
32	91327,42	153,522	4019525,128	44,012	44,554
33	91173,898	163,996	3928274,469	43,086	43,628
34	91009,902	166,274	3837182,569	42,162	42,704
35	90843,628	178,975	3746255,804	41,239	41,781
36	90664,653	185,267	3655501,664	40,319	40,861
37	90479,386	190,773	3564929,645	39,4	39,942
38	90288,613	211,371	3474545,646	38,483	39,025
39	90077,242	211,651	3384362,718	37,572	38,114
40	89865,591	237,534	3294391,302	36,659	37,201
41	89628,057	217,242	3204644,478	35,755	36,297
42	89410,815	260,252	3115125,043	34,841	35,383
43	89150,563	254,913	3025844,354	33,941	34,483
44	88895,649	271,463	2936821,248	33,037	33,579
45	88624,186	300,835	2848061,33	32,136	32,678
46	88323,351	305,217	2759587,562	31,244	31,786
47	88018,134	317,655	2671416,819	30,351	30,893
48	87700,479	366,01	2583557,513	29,459	30,001
49	87334,469	362,504	2496040,038	28,58	29,122
50	86971,965	426,353	2408886,821	27,697	28,239
51	86545,613	401,115	2322128,032	26,831	27,373
52	86144,498	477,039	2235782,977	25,954	26,496
53	85667,459	491,862	2149876,999	25,096	25,638
54	85175,596	519,396	2064455,471	24,238	24,78
55	84656,201	570,526	1979539,573	23,383	23,925
56	84085,674	614,829	1895168,635	22,539	23,081
57	83470,846	632,437	1811390,375	21,701	22,243
58	82838,409	720,143	1728235,748	20,863	21,405
59	82118,266	723,591	1645757,411	20,041	20,583
60	81394,675	931,942	1564000,941	19,215	19,757

Continuação da Tabela E.3

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
61	80462,733	824,96	1483072,237	18,432	18,974
62	79637,772	1005,373	1403021,984	17,618	18,16
63	78632,399	1088,799	1323886,898	16,836	17,378
64	77543,6	1187,833	1245798,899	16,066	16,608
65	76355,768	1350,213	1168849,215	15,308	15,85
66	75005,555	1423,785	1093168,553	14,575	15,117
67	73581,77	1527,285	1018874,891	13,847	14,389
68	72054,484	1750,464	946056,764	13,13	13,672
69	70304,02	1740,238	874877,511	12,444	12,986
70	68563,782	2294,867	805443,61	11,747	12,289
71	66268,915	2037,825	738027,261	11,137	11,679
72	64231,091	2502,121	672777,258	10,474	11,016
73	61728,969	2629,395	609797,228	9,879	10,421
74	59099,574	2878,917	549382,957	9,296	9,838
75	56220,657	3244,859	491722,841	8,746	9,288
76	52975,798	3267,44	437124,614	8,251	8,793
77	49708,358	3398,712	385782,536	7,761	8,303
78	46309,645	3753,25	337773,535	7,294	7,836
79	42556,395	3516,361	293340,515	6,893	7,435
80	39040,034	3828,928	252542,3	6,469	7,011
81	35211,107	3211,915	215416,729	6,118	6,66
82	31999,192	3510,262	181811,58	5,682	6,224
83	28488,93	3346,396	151567,519	5,32	5,862
84	25142,534	3339,96	124751,787	4,962	5,504
85	21802,575	3293,867	101279,233	4,645	5,187
86	18508,708	2965,582	81123,592	4,383	4,925
87	15543,126	2656,074	64097,675	4,124	4,666
88	12887,052	2464,883	49882,586	3,871	4,413
89	10422,168	2029,223	38227,976	3,668	4,21

Continuação da Tabela E.3					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
90	8392,945	1937,586	28820,419	3,434	3,976
91	6455,36	1360,266	21396,267	3,314	3,856
92	5095,094	1217,827	15621,04	3,066	3,608
93	3877,267	955,171	11134,86	2,872	3,414
94	2922,095	762,384	7735,179	2,647	3,189
95	2159,711	640,068	5194,276	2,405	2,947
96	1519,643	468,928	3354,598	2,207	2,749
97	1050,715	338,784	2069,42	1,97	2,512
98	711,93	263,639	1188,097	1,669	2,211
99	448,291	155,878	607,987	1,356	1,898
100	292,413	109,556	0	0	0,542
Fim da Tabela					

Tabela E.4: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sexo Feminino.

Início da Tabela					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
0	100000	2952,115	7312296,503	73,123	73,665
1	97047,885	487,153	7214776,28	74,342	74,884
2	96560,731	231,94	7117971,972	73,715	74,257
3	96328,792	143,633	7021527,211	72,891	73,433
4	96185,158	103,554	6925270,236	71,999	72,541
5	96081,604	79,214	6829136,854	71,076	71,618
6	96002,39	69,644	6733094,857	70,135	70,677
7	95932,747	58,962	6637127,289	69,185	69,727
8	95873,785	51,656	6541224,023	68,227	68,769
9	95822,129	47,154	6445376,066	67,264	67,806

Continuação da Tabela E.4

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
10	95774,975	45,426	6349577,514	66,297	66,839
11	95729,549	43,631	6253825,252	65,328	65,87
12	95685,917	43,249	6158117,519	64,358	64,9
13	95642,669	45,525	6062453,226	63,386	63,928
14	95597,143	51,134	5966833,319	62,416	62,958
15	95546,01	57,837	5871261,743	61,45	61,992
16	95488,173	62,87	5775744,651	60,486	61,028
17	95425,303	65,508	5680287,913	59,526	60,068
18	95359,795	72,611	5584895,365	58,567	59,109
19	95287,184	76,882	5489571,875	57,611	58,153
20	95210,302	82,012	5394323,132	56,657	57,199
21	95128,29	83,723	5299153,836	55,705	56,247
22	95044,567	86,199	5204067,408	54,754	55,296
23	94958,368	89,334	5109065,941	53,803	54,345
24	94869,034	93,585	5014152,24	52,853	53,395
25	94775,448	92,554	4919329,999	51,905	52,447
26	94682,894	97,15	4824600,828	50,955	51,497
27	94585,744	102,4	4729966,509	50,007	50,549
28	94483,343	103,617	4635431,965	49,061	49,603
29	94379,726	105,776	4541000,431	48,114	48,656
30	94273,95	115,207	4446673,593	47,168	47,71
31	94158,743	113,867	4352457,247	46,225	46,767
32	94044,876	122,268	4258355,438	45,28	45,822
33	93922,608	131,475	4164371,696	44,338	44,88
34	93791,133	135,491	4070514,826	43,4	43,942
35	93655,641	147,07	3976791,439	42,462	43,004
36	93508,571	152,722	3883209,332	41,528	42,07
37	93355,849	159,429	3789777,122	40,595	41,137
38	93196,42	176,625	3696500,987	39,664	40,206

Continuação da Tabela E.4

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
39	93019,795	182,974	3603392,88	38,738	39,28
40	92836,821	202,987	3510464,572	37,813	38,355
41	92633,834	194,123	3417729,244	36,895	37,437
42	92439,712	226,942	3325192,471	35,971	36,513
43	92212,77	227,606	3232866,231	35,059	35,601
44	91985,164	243,162	3140767,264	34,144	34,686
45	91742,002	269,083	3048903,681	33,233	33,775
46	91472,919	278,571	2957296,221	32,33	32,872
47	91194,348	293,196	2865962,588	31,427	31,969
48	90901,152	334,054	2774914,838	30,527	31,069
49	90567,098	335,679	2684180,713	29,637	30,179
50	90231,419	388,159	2593781,455	28,746	29,288
51	89843,259	376,336	2503744,116	27,868	28,41
52	89466,923	438,626	2414089,025	26,983	27,525
53	89028,297	456,244	2324841,416	26,114	26,656
54	88572,052	480,591	2236041,241	25,245	25,787
55	88091,461	527,571	2147709,485	24,38	24,922
56	87563,889	567,831	2059881,81	23,524	24,066
57	86996,059	589,045	1972601,836	22,675	23,217
58	86407,014	665,445	1885900,3	21,826	22,368
59	85741,569	676,538	1799826,008	20,991	21,533
60	85065,031	843,693	1714422,708	20,154	20,696
61	84221,338	774,93	1629779,524	19,351	19,893
62	83446,407	926,521	1545945,651	18,526	19,068
63	82519,886	1004,934	1462962,504	17,729	18,271
64	81514,952	1099,481	1380945,085	16,941	17,483
65	80415,471	1241,664	1299979,874	16,166	16,708
66	79173,807	1317,617	1220185,235	15,411	15,953
67	77856,19	1419,858	1141670,237	14,664	15,206

Continuação da Tabela E.4

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
68	76436,332	1621,759	1064523,976	13,927	14,469
69	74814,573	1628,148	988898,523	13,218	13,76
70	73186,425	2106,543	914898,024	12,501	13,043
71	71079,882	1923,358	842764,87	11,857	12,399
72	69156,524	2354,802	772646,667	11,172	11,714
73	66801,722	2492,515	704667,544	10,549	11,091
74	64309,206	2741,726	639112,08	9,938	10,48
75	61567,48	3101,532	576173,736	9,358	9,9
76	58465,948	3164,2	516157,022	8,828	9,37
77	55301,748	3328,819	459273,174	8,305	8,847
78	51972,929	3704,86	405635,835	7,805	8,347
79	48268,07	3512,348	355515,336	7,365	7,907
80	44755,721	3914,374	309003,44	6,904	7,446
81	40841,347	3327,63	266204,906	6,518	7,06
82	37513,718	3720,701	227027,374	6,052	6,594
83	33793,016	3593,131	191374,007	5,663	6,205
84	30199,885	3648,342	159377,556	5,277	5,819
85	26551,543	3686,937	131001,842	4,934	5,476
86	22864,606	3361,356	106293,768	4,649	5,191
87	19503,25	3078,984	85109,84	4,364	4,906
88	16424,266	2921,03	67146,082	4,088	4,63
89	13503,236	2426,397	52182,33	3,864	4,406
90	11076,84	2423,263	39892,292	3,601	4,143
91	8653,577	1684,447	30027,084	3,47	4,012
92	6969,13	1571,508	22215,731	3,188	3,73
93	5397,621	1261,589	16032,355	2,97	3,512
94	4136,033	1031,303	11265,528	2,724	3,266
95	3104,729	890,729	7645,147	2,462	3,004
96	2214,001	661,75	4985,782	2,252	2,794

Continuação da Tabela E.4					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
97	1552,251	489,1	3102,657	1,999	2,541
98	1063,151	384,97	1794,956	1,688	2,23
99	678,181	232,526	924,29	1,363	1,905
100	445,655	166,565	0	0	0,542
Fim da Tabela					

Tabela E.5: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pela Extensão do Método Lee-Carter - Sem Distinção de Sexo.

Início da Tabela					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
0	100000	5062,494	6667713,816	66,677	67,219
1	94937,506	837,331	6570245,063	69,206	69,748
2	94100,176	379,019	6475726,222	68,817	69,359
3	93721,157	221,304	6381815,556	68,094	68,636
4	93499,852	161,077	6288205,052	67,254	67,796
5	93338,776	121,841	6194785,738	66,369	66,911
6	93216,935	102,686	6101507,882	65,455	65,997
7	93114,249	92,831	6008342,291	64,527	65,069
8	93021,418	78,954	5915274,457	63,59	64,132
9	92942,463	72,954	5822292,517	62,644	63,186
10	92869,509	68,658	5729386,531	61,693	62,235
11	92800,851	67,225	5636551,351	60,738	61,28
12	92733,626	66,376	5543784,113	59,782	60,324
13	92667,25	70,826	5451083,675	58,824	59,366
14	92596,424	81,367	5358451,838	57,869	58,411
15	92515,056	93,781	5265896,098	56,919	57,461
16	92421,275	112,988	5173427,933	55,977	56,519

Continuação da Tabela E.5

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
17	92308,287	127,402	5081063,151	55,044	55,586
18	92180,884	147,468	4988818,566	54,12	54,662
19	92033,416	157,756	4896711,416	53,206	53,748
20	91875,66	165,634	4804756,877	52,296	52,838
21	91710,026	169,836	4712964,034	51,39	51,932
22	91540,19	174,97	4621338,926	50,484	51,026
23	91365,22	176,708	4529886,221	49,58	50,122
24	91188,512	181,779	4438609,355	48,675	49,217
25	91006,733	180,129	4347511,733	47,771	48,313
26	90826,603	183,945	4256595,065	46,865	47,407
27	90642,659	187,963	4165860,435	45,959	46,501
28	90454,696	194,608	4075311,757	45,054	45,596
29	90260,088	189,403	3984954,366	44,15	44,692
30	90070,685	203,557	3894788,979	43,241	43,783
31	89867,128	196,162	3804820,073	42,338	42,88
32	89670,966	213,996	3715051,026	41,43	41,972
33	89456,97	225,22	3625487,058	40,528	41,07
34	89231,75	228,862	3536142,699	39,629	40,171
35	89002,888	244,308	3447025,38	38,729	39,271
36	88758,58	253,342	3358144,646	37,835	38,377
37	88505,238	262,182	3269512,737	36,941	37,483
38	88243,055	285,502	3181138,591	36,05	36,592
39	87957,553	286,691	3093038,287	35,165	35,707
40	87670,861	319,959	3005224,08	34,278	34,82
41	87350,902	302,201	2917713,198	33,402	33,944
42	87048,701	356,847	2830513,397	32,516	33,058
43	86691,854	355,796	2743643,119	31,648	32,19
44	86336,057	382,597	2657129,164	30,777	31,319
45	85953,46	419,354	2570984,405	29,911	30,453

Continuação da Tabela E.5

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
46	85534,106	422,839	2485240,622	29,056	29,598
47	85111,267	445,682	2399917,935	28,197	28,739
48	84665,585	504,646	2315029,509	27,343	27,885
49	84160,939	512,393	2230616,247	26,504	27,046
50	83648,546	584,939	2146711,505	25,663	26,205
51	83063,607	554,236	2063355,429	24,841	25,383
52	82509,371	648,081	1980568,94	24,004	24,546
53	81861,29	679,5	1898383,609	23,19	23,732
54	81181,791	720,447	1816862,069	22,38	22,922
55	80461,343	781,882	1736040,502	21,576	22,118
56	79679,461	835,665	1655970,099	20,783	21,325
57	78843,796	867,419	1576708,471	19,998	20,54
58	77976,377	964,006	1498298,385	19,215	19,757
59	77012,371	986,088	1420804,011	18,449	18,991
60	76026,283	1191,258	1344284,684	17,682	18,224
61	74835,025	1097,444	1268854,031	16,955	17,497
62	73737,581	1293,309	1194567,728	16,2	16,742
63	72444,272	1389,197	1121476,801	15,481	16,023
64	71055,075	1485,481	1049727,128	14,773	15,315
65	69569,594	1639,643	979414,793	14,078	14,62
66	67929,951	1723,543	910665,021	13,406	13,948
67	66206,408	1818,861	843596,841	12,742	13,284
68	64387,547	2031,258	778299,863	12,088	12,63
69	62356,29	2018,408	714927,945	11,465	12,007
70	60337,882	2494,788	653580,859	10,832	11,374
71	57843,094	2261,435	594490,371	10,278	10,82
72	55581,66	2672,782	537777,994	9,675	10,217
73	52908,878	2756,777	483532,725	9,139	9,681
74	50152,1	2934,088	432002,236	8,614	9,156

Continuação da Tabela E.5

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
75	47218,013	3189,381	383317,179	8,118	8,66
76	44028,632	3173,564	337693,857	7,67	8,212
77	40855,068	3260,598	295252,007	7,227	7,769
78	37594,47	3474,625	256027,239	6,81	7,352
79	34119,845	3216,008	220170,082	6,453	6,995
80	30903,837	3353,657	187658,241	6,072	6,614
81	27550,18	2840,875	158431,232	5,751	6,293
82	24709,305	3002,429	132301,49	5,354	5,896
83	21706,875	2819,612	109093,4	5,026	5,568
84	18887,263	2746,858	88796,33	4,701	5,243
85	16140,405	2622,998	71282,496	4,416	4,958
86	13517,407	2317,987	56453,59	4,176	4,718
87	11199,419	2050,006	44095,177	3,937	4,479
88	9149,414	1862,272	33920,761	3,707	4,249
89	7287,141	1502,7	25702,483	3,527	4,069
90	5784,441	1385,979	19166,692	3,313	3,855
91	4398,462	975,059	14075,24	3,2	3,742
92	3423,403	856,984	10164,308	2,969	3,511
93	2566,42	661,314	7169,396	2,794	3,336
94	1905,105	513,908	4933,634	2,59	3,132
95	1391,197	420,277	3285,483	2,362	2,904
96	970,92	307,629	2104,424	2,167	2,709
97	663,291	218,037	1287,319	1,941	2,483
98	445,254	168,619	733,047	1,646	2,188
99	276,635	98,195	372,102	1,345	1,887
100	178,44	67,751	0	0	0,542

Fim da Tabela

Tabela E.6: Cálculo das taxas vitalícias de 1970 obtidas a partir da tábua de mortalidade de 1970 elaborada através da taxa de mortalidade ajustada pelo Método Lee-Carter - Sem Distinção de Sexo.

Início da Tabela					
Idade	$\dot{l}_{x,1970}$	$\dot{d}_{x,1970}$	$\dot{T}_{x,1970}$	$\dot{e}_{x,1970}$	Taxas Vitalícias
0	100000	3413,039	6957026,355	69,57	70,112
1	96586,961	538,435	6858732,875	71,011	71,553
2	96048,526	266,554	6762415,131	70,406	70,948
3	95781,973	162,857	6666499,882	69,601	70,143
4	95619,116	120,343	6570799,337	68,718	69,26
5	95498,773	93,253	6475240,393	67,804	68,346
6	95405,52	78,757	6379788,247	66,87	67,412
7	95326,763	72,35	6284422,106	65,925	66,467
8	95254,413	62,234	6189131,518	64,975	65,517
9	95192,179	57,6	6093908,222	64,017	64,559
10	95134,579	55,1	5998744,843	63,055	63,597
11	95079,478	54,255	5903637,815	62,092	62,634
12	95025,223	53,411	5808585,464	61,127	61,669
13	94971,812	56,49	5713586,947	60,161	60,703
14	94915,322	65,711	5618643,38	59,196	59,738
15	94849,611	77,106	5523760,913	58,237	58,779
16	94772,504	92,42	5428949,855	57,284	57,826
17	94680,084	103,39	5334223,561	56,339	56,881
18	94576,694	120,124	5239595,172	55,4	55,942
19	94456,57	129,434	5145078,54	54,47	55,012
20	94327,136	135,772	5050686,687	53,544	54,086
21	94191,364	139,168	4956427,438	52,621	53,163
22	94052,196	140,857	4862305,658	51,698	52,24
23	93911,338	143,501	4768323,891	50,775	51,317
24	93767,838	148,97	4674484,303	49,852	50,394
25	93618,868	148,789	4580790,95	48,93	49,472

Continuação da Tabela E.6

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
26	93470,079	152,776	4487246,477	48,007	48,549
27	93317,303	156,773	4393852,785	47,085	47,627
28	93160,53	162,531	4300613,869	46,163	46,705
29	92998	161,412	4207534,604	45,243	45,785
30	92836,587	172,442	4114617,31	44,321	44,863
31	92664,145	170,3	4021866,944	43,403	43,945
32	92493,846	183,287	3929287,948	42,482	43,024
33	92310,559	193,413	3836885,746	41,565	42,107
34	92117,145	199,694	3744671,894	40,651	41,193
35	91917,452	213,333	3652654,596	39,738	40,28
36	91704,118	221,565	3560843,811	38,83	39,372
37	91482,553	231,793	3469250,475	37,923	38,465
38	91250,76	252,161	3377883,819	37,018	37,56
39	90998,598	259,887	3286759,14	36,119	36,661
40	90738,711	287,172	3195890,485	35,221	35,763
41	90451,539	281,241	3105295,36	34,331	34,873
42	90170,298	324,963	3014984,441	33,437	33,979
43	89845,335	331,12	2924976,624	32,556	33,098
44	89514,215	355,888	2835296,849	31,674	32,216
45	89158,327	389,295	2745960,578	30,799	31,341
46	88769,032	400,283	2656996,898	29,932	30,474
47	88368,749	424,731	2568428,008	29,065	29,607
48	87944,018	477,378	2480271,624	28,203	28,745
49	87466,64	490,458	2392566,294	27,354	27,896
50	86976,182	553,864	2305344,883	26,505	27,047
51	86422,319	537,485	2218645,632	25,672	26,214
52	85884,834	619,75	2132492,056	24,83	25,372
53	85265,084	654,261	2046917,097	24,007	24,549
54	84610,823	691,101	1961979,143	23,188	23,73

Continuação da Tabela E.6

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
55	83919,722	750,645	1877713,87	22,375	22,917
56	83169,078	800,888	1794169,47	21,573	22,115
57	82368,19	836,948	1711400,836	20,777	21,319
58	81531,242	925,361	1629451,12	19,986	20,528
59	80605,882	951,362	1548382,558	19,209	19,751
60	79654,519	1131,048	1468252,358	18,433	18,975
61	78523,471	1062,889	1389163,362	17,691	18,233
62	77460,582	1240,757	1311171,336	16,927	17,469
63	76219,825	1331,962	1234331,132	16,194	16,736
64	74887,863	1425,365	1158777,288	15,473	16,015
65	73462,498	1571,303	1084602,107	14,764	15,306
66	71891,195	1654,937	1011925,26	14,076	14,618
67	70236,258	1752,454	940861,534	13,396	13,938
68	68483,804	1953,058	871501,503	12,726	13,268
69	66530,746	1949,093	803994,227	12,085	12,627
70	64581,653	2391,02	738438,028	11,434	11,976
71	62190,633	2198,506	675051,885	10,855	11,397
72	59992,128	2601,571	613960,504	10,234	10,776
73	57390,557	2699,432	555269,162	9,675	10,217
74	54691,125	2883,604	499228,321	9,128	9,67
75	51807,522	3154,761	445978,998	8,608	9,15
76	48652,761	3162,593	395748,856	8,134	8,676
77	45490,168	3280,405	348677,391	7,665	8,207
78	42209,763	3525,934	304827,426	7,222	7,764
79	38683,829	3284,394	264380,63	6,834	7,376
80	35399,435	3513,146	227338,998	6,422	6,964
81	31886,289	2990,596	193696,136	6,075	6,617
82	28895,693	3235,569	163305,146	5,652	6,194
83	25660,123	3071,739	136027,238	5,301	5,843

Continuação da Tabela E.6

Idade	$i_{x,1970}$	$d_{x,1970}$	$T_{x,1970}$	$e_{x,1970}$	Taxas Vitalícias
84	22588,385	3043,687	111902,984	4,954	5,496
85	19544,698	2968,487	90836,443	4,648	5,19
86	16576,211	2655,907	72775,988	4,39	4,932
87	13920,304	2384,513	57527,731	4,133	4,675
88	11535,791	2215,939	44799,684	3,884	4,426
89	9319,852	1800,938	34371,862	3,688	4,23
90	7518,914	1730,356	25952,479	3,452	3,994
91	5788,558	1194,868	19298,743	3,334	3,876
92	4593,691	1094,569	14107,618	3,071	3,613
93	3499,122	860,868	10061,212	2,875	3,417
94	2638,254	685,811	6992,524	2,65	3,192
95	1952,443	575,127	4697,175	2,406	2,948
96	1377,317	425,854	3032,295	2,202	2,744
97	951,462	307,104	1867,906	1,963	2,505
98	644,358	239,634	1069,996	1,661	2,203
99	404,724	142,54	545,455	1,348	1,89
100	262,183	100,363	0	0	0,542

Fim da Tabela

Apêndice F

Análise Descritiva dos Dados da Companhia

Tabela F.1: Medidas descritivas dos dados de exposição ao risco dos três casos de estudo.

Medida Descritiva	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
início	2007	2007	2007
fim	2021	2021	2021
n° de observações	1515	1515	1515
mínimo	0	0	0
1° quartil	8,750	6,000	15,500
média	38,275	18,854	56,675
mediana	26,500	15,500	43,000
3° quartil	63,500	29,500	92,000
máximo	159	71	215
desvio padrão	34,671	14,944	47,959

Tabela F.2: Medidas descritivas dos dados de número de mortes dos três casos de estudo.

Medida Descritiva	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
início	2007	2007	2007
fim	2021	2021	2021
n° de observações	1515	1515	1515
mínimo	0	0	0
1° quartil	0	0	0
média	0,743	0,268	1,011
mediana	0	0	0
3° quartil	1,000	0,000	1,500
máximo	6	5	8
desvio padrão	1,231	0,614	1,523

Tabela F.3: Médias anuais dos dados de exposição ao risco dos três casos de estudo.

Ano	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
2007	43,550	18,366	55,109
2008	43,550	19,926	63,475
2009	45,569	20,946	66,515
2010	45,975	21,218	67,193
2011	44,436	21,381	65,817
2012	39,906	20,951	60,856
2013	39,495	19,762	59,257
2014	39,906	20,951	60,856
2015	37,708	19,500	57,208
2016	34,951	18,342	53,292
2017	31,777	16,589	48,366
2018	29,342	16,015	45,356
2019	31,941	16,208	48,149
2020	33,817	16,495	50,312
2021	32,198	16,168	48,366

Tabela F.4: Médias anuais dos dados de número de mortes dos três casos de estudo.

Ano	Sexo Masculino	Sexo Feminino	Sem Distinção de Sexo
2007	0,396	0,129	0,525
2008	0,257	0,069	0,327
2009	1,079	0,347	1,426
2010	0,673	0,238	0,911
2011	0,871	0,257	1,129
2012	0,851	0,307	1,158
2013	0,713	0,267	0,980
2014	0,851	0,307	91,158
2015	0,713	0,287	1,000
2016	0,861	0,396	1,257
2017	0,871	0,416	1,287
2018	0,941	0,287	1,228
2019	0,931	0,287	1,218
2020	0,564	0,198	0,762
2021	0,564	0,228	0,792

Apêndice G

Comparação das taxas de mortalidade da população portuguesa e da população da companhia

Tabela G.1: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Masculino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000247	0,015730	0,392067	0,008567	1,003759
	18-30 anos	0,000109	0,010430	0,752090	0,008384	1,774980
	30-40 anos	0,000006	0,002332	0,643140	0,001890	0,490475
	40-50 anos	0,000003	0,001818	0,227763	0,001338	0,773227
	50-60 anos	0,000016	0,003997	0,273840	0,003771	1,871212
	60-65 anos	0,000015	0,003853	0,175919	0,002625	1,915523
Sexo Masculino	65-70 anos	0,000057	0,007520	0,420621	0,006630	2,000966
	70-75 anos	0,000013	0,003616	0,066360	0,003091	0,659588
	75-80 anos	0,000325	0,018037	0,250353	0,015202	6,931709
	80-85 anos	0,001216	0,034866	0,438275	0,033266	2,350231
	85-90 anos	0,001308	0,036170	0,134009	0,023205	1,510266

Tabela G.2: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Masculino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000273	0,016533	0,396988	0,008936	1,017455
	18-30 anos	0,000110	0,010476	0,761184	0,008440	1,789430
	30-40 anos	0,000005	0,002334	0,643030	0,001901	0,492753
	40-50 anos	0,000003	0,001780	0,219863	0,001287	0,750470
	50-60 anos	0,000015	0,003831	0,261359	0,003601	1,801996
	60-65 anos	0,000016	0,003998	0,179604	0,002671	2,023110
Sexo Masculino	65-70 anos	0,000063	0,007945	0,446728	0,007120	2,110034
	70-75 anos	0,000012	0,003505	0,067288	0,003110	0,615175
	75-80 anos	0,000380	0,019487	0,277040	0,016784	7,492006
	80-85 anos	0,001374	0,037061	0,466094	0,035586	2,490521
	85-90 anos	0,001438	0,037917	0,142696	0,024530	1,576672

Tabela G.3: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Masculino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000248	0,015751	0,542057	0,009227	0,998005
	18-30 anos	0,000139	0,011790	0,744572	0,009130	1,846112
	30-40 anos	0,000006	0,002518	1,575299	0,002170	0,881269
	40-50 anos	0,000003	0,001764	0,210241	0,001270	0,985434
	50-60 anos	0,000019	0,004387	0,304039	0,004244	2,859921
	60-65 anos	0,000006	0,002426	0,097248	0,001664	1,033402
Sexo Masculino	65-70 anos	0,000083	0,009115	0,564837	0,007789	1,688365
	70-75 anos	0,000010	0,003194	0,063438	0,002765	0,897088
	75-80 anos	0,000336	0,018340	0,237443	0,014406	6,391340
	80-85 anos	0,001447	0,038036	0,483247	0,035030	2,197785
	85-90 anos	0,000952	0,030861	0,138956	0,026443	1,255847

Tabela G.4: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Masculino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000272	0,016496	0,542352	0,009522	0,988226
	18-30 anos	0,000140	0,011834	0,748463	0,009178	1,863303
	30-40 anos	0,000006	0,002522	1,553968	0,002171	0,862414
	40-50 anos	0,000003	0,001729	0,208558	0,001248	0,967550
	50-60 anos	0,000018	0,004219	0,292056	0,004074	2,759152
	60-65 anos	0,000006	0,002474	0,092596	0,001578	1,096373
Sexo Masculino	65-70 anos	0,000089	0,009454	0,587546	0,008062	1,756527
	70-75 anos	0,000007	0,002642	0,051035	0,002245	0,752798
	75-80 anos	0,000389	0,019729	0,263869	0,015988	6,870690
	80-85 anos	0,001614	0,040174	0,512017	0,037350	2,316113
	85-90 anos	0,001047	0,032355	0,145523	0,027516	1,301152

Tabela G.5: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Feminino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000199	0,014096	Inf	0,008965	NaN
	18-30 anos	0,000100	0,010007	0,948618	0,009364	2,690543
	30-40 anos	0,000061	0,007829	0,874618	0,007110	2,169512
	40-50 anos	0,000012	0,003441	Inf	0,002419	NaN
	50-60 anos	0,000001	0,001006	0,252498	0,000916	0,708957
	60-65 anos	0,000005	0,002212	0,436361	0,002098	1,176072
Sexo Feminino	65-70 anos	0,000013	0,003599	0,254452	0,002681	2,041712
	70-75 anos	0,000013	0,003619	0,170713	0,002895	0,358033
	75-80 anos	0,000363	0,019047	0,605052	0,017344	3,456826
	80-85 anos	0,001238	0,035188	0,693696	0,034319	5,122604
	85-90 anos	0,000869	0,029476	0,270329	0,025824	1,297799

Tabela G.6: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Feminino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000223	0,014917	Inf	0,009420	NaN
	18-30 anos	0,000098	0,009916	0,936436	0,009267	2,664888
	30-40 anos	0,000060	0,007775	0,862722	0,007046	2,148885
	40-50 anos	0,000012	0,003414	Inf	0,002403	NaN
	50-60 anos	0,000001	0,001051	0,257794	0,000921	0,728775
	60-65 anos	0,000005	0,002281	0,454720	0,002134	1,235020
Sexo Feminino	65-70 anos	0,000015	0,003854	0,277320	0,002929	2,212659
	70-75 anos	0,000016	0,004023	0,187061	0,003146	0,412913
	75-80 anos	0,000406	0,020141	0,644224	0,018454	3,652766
	80-85 anos	0,001378	0,037118	0,732213	0,036215	5,413844
	85-90 anos	0,001019	0,031925	0,296695	0,028882	1,424662

Tabela G.7: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através da Extensão do Método Lee-Carter - Sexo Feminino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000194	0,013944	Inf	0,008936	NaN
	18-30 anos	0,000131	0,011439	0,933486	0,010548	2,891713
	30-40 anos	0,000073	0,008559	0,853271	0,007506	2,429747
	40-50 anos	0,000014	0,003789	Inf	0,002702	NaN
	50-60 anos	0,000001	0,001104	0,212290	0,000874	0,936320
	60-65 anos	0,000008	0,002857	0,634706	0,002698	1,584932
Sexo Feminino	65-70 anos	0,000009	0,002980	0,238927	0,002863	1,639647
	70-75 anos	0,000006	0,002373	0,108703	0,002192	0,568976
	75-80 anos	0,000310	0,017610	0,538140	0,015799	3,554616
	80-85 anos	0,001371	0,037022	0,753723	0,036071	6,414511
	85-90 anos	0,000676	0,026010	0,231510	0,020345	1,031783

Tabela G.8: Medidas de Avaliação resultantes da comparação entre a média das taxas de mortalidade da companhia entre os anos 2015-2018, obtida através de Médias Móveis Ponderadas, e a taxa de mortalidade da população portuguesa do ano 1997, obtida através do Método Lee-Carter - Sexo Feminino.

	Faixa etária	EQM	REQM	EPAM	EAM	U-Theil
	18-90 anos	0,000215	0,014651	Inf	0,00928	NaN
	18-30 anos	0,000129	0,011350	0,919074	0,010451	2,866527
	30-40 anos	0,000072	0,008507	0,837652	0,007442	2,395876
	40-50 anos	0,000014	0,003762	Inf	0,002699	NaN
	50-60 anos	0,000001	0,001143	0,222064	0,000895	0,973950
	60-65 anos	0,000008	0,002904	0,657781	0,002734	1,652342
Sexo Feminino	65-70 anos	0,000010	0,003156	0,253123	0,003022	1,791721
	70-75 anos	0,000006	0,002476	0,116763	0,002250	0,606830
	75-80 anos	0,000349	0,018673	0,572150	0,016801	3,770401
	80-85 anos	0,001518	0,038956	0,793712	0,037968	6,756017
	85-90 anos	0,000783	0,027989	0,250597	0,022297	1,133465