

UNIVERSIDADE FEDERAL DO PARANÁ

GUERINO PIROLLO JUNIOR

GRADUAÇÃO BAYESIANA E PROJEÇÃO DE TAXAS DE  
MORTALIDADE APLICADAS À POPULAÇÕES DE FUNDOS DE  
PENSÃO DO SETOR ELÉTRICO

CURITIBA  
2009

GUERINO PIROLLO JUNIOR

GRADUAÇÃO BAYESIANA E PROJEÇÃO DE TAXAS DE MORTALIDADE  
APLICADAS À POPULAÇÕES DE FUNDOS DE PENSÃO DO SETOR ELÉTRICO

Dissertação apresentada ao Curso de Pós-Graduação em Métodos Numéricos em Engenharia do Setor de Tecnologia do Centro de Estudos em Engenharia Civil Professor Inaldo Ayres Vieira da Universidade Federal do Paraná, como parte das exigências para obtenção do título de Mestre em Ciências.

Orientador: Prof. PhD. Paulo Justiniano Ribeiro Junior

CURITIBA

2009

## **TERMO DE APROVAÇÃO**

**GUERINO PIROLLO JUNIOR**

**GRADUAÇÃO BAYESIANA E PROJEÇÃO DE TAXAS DE MORTALIDADE  
APLICADAS À POPULAÇÕES DE FUNDOS DE PENSÃO DO SETOR ELÉTRICO**

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre no curso de Pós-Graduação em Métodos Numéricos em Engenharia do Setor de Tecnologia do Centro de Estudos em Engenharia Civil Professor Inaldo Ayres Vieira, da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientador: Prof. PhD. Paulo Justiniano Ribeiro Junior  
Universidade Federal do Paraná

Prof. PhD. Kaizô Iwakami Beltrão  
Escola Nacional de Ciências Estatísticas.

Prof. PhD. Renato Martins Assunção  
Universidade Federal de Minas Gerais

Curitiba, 21 de dezembro de 2009.

À Elaine, minha esposa, por ter suportado as ausências e faltas impostas por este trabalho.

À Joana, minha mãe, por sempre ter acreditado em mim e por suas constantes orações.

## AGRADECIMENTOS

Ao Pai celestial, pela vida, pelos dons e portas que sempre me abriu.

Ao professor Paulo Justiniano, por sua disposição em compartilhar seu grande conhecimento em Estatística, por sua incrível capacidade em tornar assuntos tão complexos em linguagens acessíveis e, principalmente, por sua orientação.

Ao professor Jorge Festa, pela confiança, amizade e pelos anos de convívio na Fundação Copel, que muito contribuiu em minha trajetória estudantil e profissional.

Aos colegas atuários Carmem, Luciano e Silvio, por terem intermediado junto às suas Fundações, o fornecimento de dados para a realização deste trabalho.

Aos colegas de pós-graduação Ana Beatriz, Helena e Reginaldo, pela amizade, companheirismo e colaboração recebida durante o curso.

À Maristela, por sua prestatividade, competência e eficiência, sempre simpática.

À todos os professores do PPGMNE, que também possuem sua parcela de contribuição à este trabalho através dos conhecimentos repassados.

À Fundação Copel, por me disponibilizar o tempo necessário ao estudo, entendendo a relevância deste trabalho também para a entidade.

## RESUMO

O processo de envelhecimento populacional que vem afetando a população mundial - o qual é caracterizado por quedas significativas nas taxas de mortalidade e fecundidade - tem afetado os sistemas de previdência e seguridade em todo o mundo, alertando-os a adotarem hipóteses biométricas cada vez mais conservadoras, impactando os cálculos atuariais. Essas tendências já começam a ser contempladas nos cálculos de provisões matemáticas e de benefícios através do uso de algumas ferramentas de construção de tábuas de mortalidade e projeções de taxas de mortalidade futuras. Diante disso, considerando a população de expostos ao risco de seis Entidades de Previdência Complementar Fechadas do sul do país – especificamente do setor elétrico – o objetivo deste estudo é graduar e projetar as taxas de mortalidade desta população, obtendo informações sobre seus níveis de mortalidade atuais e futuros, auxiliando na adoção da hipótese de mortalidade geral que melhor reflita a mortalidade do grupo. Para isto, aplicou-se a técnica de Graduação Bayesiana de Taxas de Mortalidade utilizando simulação estocástica de Monte Carlo via Cadeias de Markov (MCMC) na população de assistidos das referidas Fundações, considerando os anos de estudo de 2000 a 2008, obtendo-se uma tábua de mortalidade para as idades entre 42 e 90 anos. Também foi utilizado o modelo demográfico de Lee-Carter para projeção das taxas de mortalidade graduadas, obtendo os respectivos fatores de melhoria a serem aplicados às mesmas, projetando-as para o futuro, tornando a tábua na forma geracional e obtendo o comportamento futuro dos níveis de mortalidade. A componente  $k_t$  do modelo de Lee-Carter, que corresponde a uma série temporal que descreve a mudança na mortalidade ao longo do tempo, foi modelada por um processo ARIMA, o qual foi simulado estocasticamente utilizando *bootstrap* paramétrico, obtendo-se intervalos de confiança empíricos para as taxas de mortalidade observadas.

Palavras-chave: Graduação Bayesiana. Tabua de Mortalidade. MCMC. Lee-Carter. Taxas de Mortalidade.

## ABSTRACT

The aging process which comes affecting the global population - which is characterized by the significant decreases on the mortality and fertility rates - is affecting the welfare and security systems all over the world, warning them to adopt more conservatives biometric hypotheses which impact the actuarial calculation. These trends, already begun to be considered in the mathematical provisions and benefits calculation through the use of some tools like the construction of mortality tables and future mortality rates projections. Whereupon, considering the population exposed to risk of six Pension Funds of country's southern region - specifically of the electric sector - the objective of this study is to graduate and project the mortality rates of this population to obtain information about the present and future mortality levels, which will help on the adoption of general mortality hypothesis that better reflects the mortality of the group. For that, it was applied the Bayesian Graduation Mortality Rate using stochastic simulation of Monte Carlo Markov Chain (MCMC) on the retired population of the referred funds, considering the years 2000 to 2008, to obtain a mortality table for the ages between 42 and 90 years old. It was also use the demographic Lee-Carter model for the mortality graduated rates projection, getting the improvement factors to be applied, projecting them to the future, transforming the table in the generational form and obtaining the future behavior of the mortality levels. The  $k_t$  component of the Lee-Carter model, that correspond to a time series that describes the change in the mortality over the time, was modeled by an ARIMA process, which was stochastically simulated using parametric bootstrap, getting empiric confidence intervals for the observed mortality rates.

Key-words: Bayesian Graduation. Mortality Table. MCMC. Lee-Carter. Mortality Rates.

## LISTA DE SIGLAS

ARIMA	- <i>Autorregressive Integrated Moving Average</i>
BUGS	- <i>Bayesian Inference Using Gibbs Sampler</i>
CELOS	- Fundação Celesc de Seguridade Social
CGPC	- Conselho de Gestão de Previdência Complementar
DIC	- <i>Deviance Information Criterion</i>
EFPC	- Entidade Fechada de Previdência Complementar
ELOS	- Fundação Eletrosul de Previdência e Assistência Social
FC	- Fundação Copel de Previdência e Assistência Social
FIBRA	- Fundação Itaipu Brasil de Previdência e Assistência Social
FUNCESP	- Fundação CESP de Seguridade Social
HMD	- <i>Humam Mortality Database</i>
JAGS	- <i>Just Another Gibbs Sampler</i>
MCMC	- <i>Monte Carlo Markov Chain</i>
MH	- <i>Metropolis Hastings</i>
MLG	- Modelo Linear Generalizado
SOA	- <i>Society of Actuaries</i>
SUSEP	- Superintendência de Seguros Privados
SVD	- <i>Singular Value Decomposition</i>



## SUMÁRIO

1	INTRODUÇÃO .....	11
2	METODOLOGIAS.....	17
2.1	CONCEITOS BÁSICOS DE MATEMÁTICA ATUARIAL .....	17
2.2	GRADUAÇÃO DE TAXAS DE MORTALIDADE .....	21
2.2.1	MODELOS BAYESIANOS DE GRADUAÇÃO .....	22
2.2.2	IMPLEMENTAÇÃO DOS MODELOS DE GRADUAÇÃO BAYESIANA .....	24
2.3	O MODELO DE LEE-CARTER.....	26
2.3.1	AJUSTE DO MODELO .....	29
2.3.2	AJUSTE DA COMPONENTE $k_t$ DO MODELO .....	30
2.3.3	SIMULAÇÃO DA COMPONENTE $k_t$ E TAXAS BRUTAS DE MORTALIDADE .....	30
3	GRADUAÇÃO DAS TAXAS BRUTAS DE MORTALIDADE .....	32
3.1	ANÁLISE DOS DADOS.....	33
3.2	CÁLCULO DAS TAXAS BRUTAS DE MORTALIDADE .....	38
3.3	TESTE DE ADERÊNCIA .....	39
3.4	GRADUAÇÃO BAYESIANA DAS TAXAS BRUTAS DE MORTALIDADE.....	42
3.4.1	MODELOS DE GRADUAÇÃO BAYESIANA .....	43
3.4.1.1	MODELO DE CARLIN MODIFICADO COM $\beta$ FIXO.....	45
3.4.1.2	MODELO COM REPARAMETRIZAÇÃO TIPO BROFFIT .....	51
3.4.1.3	MODELO COM FORÇA DE MORTALIDADE TIPO GOMPertz.....	55
3.4.1.4	MODELO COM FORÇA DE MORTALIDADE TIPO MAKEHAM .....	58
3.4.2	COMPARAÇÃO ENTRE OS MODELOS DE GRADUAÇÃO PROPOSTOS .....	61
3.4.3	CONSTRUÇÃO DA TÁBUA DE MORTALIDADE BAYESIANA .....	64
3.4.3.1	MODELO ESTÁTICO LOCAL .....	65
3.4.3.2	MODELO ESTÁTICO GLOBAL.....	71
3.4.4	ANÁLISE DOS RESULTADOS .....	75
4	PROJEÇÃO DE TAXAS DE MORTALIDADE .....	76
4.1	TÁBUAS ESTÁTICAS E GERACIONAIS .....	77
4.2	APLICAÇÃO DO MODELO DE LEE-CARTER.....	78
4.2.1	ESTIMAÇÃO DOS PARÂMETROS DO MODELO .....	79
4.2.2	AJUSTE DE UM PROCESSO ARIMA À COMPONENTE $k_t$ DO MODELO .....	82
4.3	SIMULAÇÃO DO PROCESSO ARIMA(0,1,1) POR <i>BOOTSTRAP</i> .....	87
4.4	PROJEÇÃO DAS TAXAS DE MORTALIDADE DA TÁBUA BAYESIANA .....	93

4.4.1	IMPACTOS NAS RESERVAS MATEMÁTICAS .....	94
5	CONCLUSÃO .....	99
	REFERÊNCIAS.....	101
	ANEXO I – Quantidade Inicial de Expostos ao Risco .....	104
	ANEXO II – Quantidade de Mortes Observadas.....	105
	ANEXO III – Taxas Brutas de Mortalidade .....	106
	ANEXO IV – Tábua de Mortalidade Bayesiana Construída sob o Modelo Estático Local.....	108
	ANEXO V – Estimativas dos Parâmetros do Modelo de Lee-Carter .....	110
	APÊNDICE .....	111

## 1 INTRODUÇÃO

Ao longo de toda a história, pessoas com 65 anos de idade ou mais nunca representaram mais do que 2% ou 3% da população. Porém nos últimos 150 anos essa proporção subiu e já alcança patamares em torno de 15% no mundo desenvolvido de hoje, se esperando um nível de 25% em 2030, podendo ser maior ainda em alguns países da Europa Continental que envelhecem rapidamente (WATSON WYATT, 1999). Este rápido envelhecimento populacional, primeiramente foi evidenciado em países desenvolvidos, mas já é realidade em países em desenvolvimento como o Brasil, que atravessa o que se convencionou chamar de terceira fase da transição demográfica (BELTRÃO; CAMARANO, 1999). Tal transição é caracterizada por uma queda nos níveis de mortalidade, seguida depois de um lapso por uma queda nos níveis de fecundidade, resultando num outro período de estabilidade.

A maior parte dos países europeus levou quase um século para completar sua transição da fecundidade. Suécia e Inglaterra, por exemplo, levaram cerca de seis décadas (aproximadamente de 1870 a 1930) para diminuir em torno de 50% seus níveis de fecundidade. O Brasil, por sua vez, experimentou um declínio similar em um quarto de século (WONG; CARVALHO, 2006). Portanto, a Transição de Estrutura Etária que o Brasil está experimentando em todas as suas regiões, tem sido extremamente rápida.

A Figura 1.1 representa as evoluções demográficas das populações da América Latina, Caribe e Brasil, segundo informações das Nações Unidas.

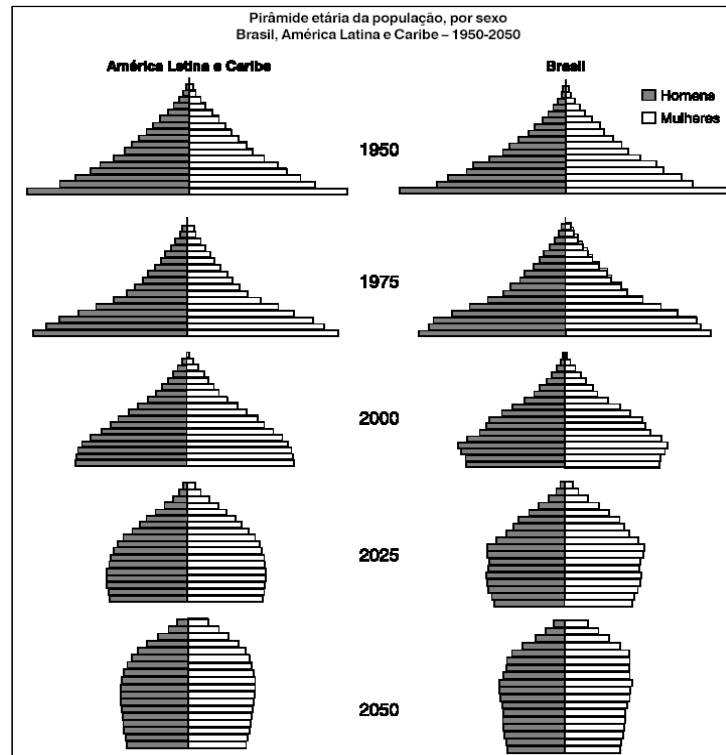


Figura 1.1 – Evolução Demográfica da América Latina, Caribe e Brasil, considerando o período compreendido entre os anos de 1950 a 2050.

O declínio da mortalidade no Brasil está se concentrando nas idades mais avançadas, tendo como efeito a aceleração do processo de envelhecimento, contribuindo também para outras duas alterações da função de sobrevivência: “retangularização” e “expansão”. A primeira diz respeito ao fato de estar ocorrendo uma grande concentração de mortes em torno de uma idade média de morte, a partir da qual a linha da função começa a se curvar. Até chegar neste ponto médio, as probabilidades de sobrevivência também vão aumentando, fazendo com que um maior número de pessoas alcance idades mais avançadas. Já a segunda alteração, é caracterizada pela elevação da idade limite que a população alcança. A Figura 2.1 ilustra a evolução da sobrevivência, caracterizada pela retangularização e expansão da função de sobrevivência  $S(x)$ , que representa a proporção de sobreviventes segundo a idade (SANTOS, 2007).

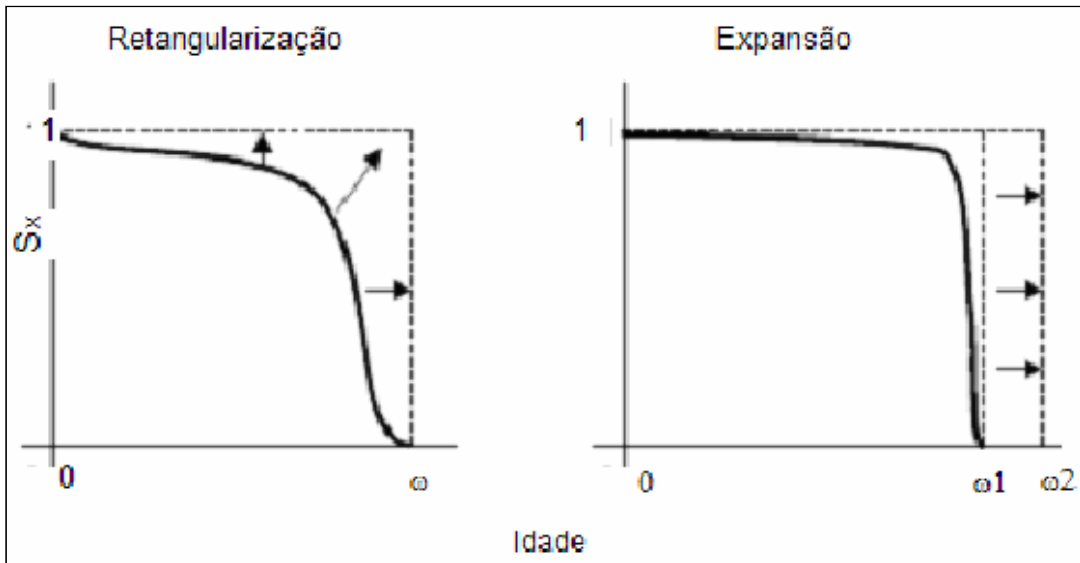


Figura 2.1 – O gráfico esquerdo representa a evolução da idade na qual ocorre o maior número de óbitos, enquanto o direito representa a evolução da idade limite a ser alcançada.

Com base na velocidade que o envelhecimento populacional vem ocorrendo no Brasil, surgem grandes desafios não só para o governo, que deverá rever e reformular suas políticas sócio-econômicas e de desenvolvimento, mas também para pesquisadores – demógrafos, estatísticos, atuários dentre outros – no sentido de dedicarem esforços na busca de soluções para o problema, propondo metodologias de adequação ao fenômeno nas mais diversas áreas, sendo uma delas previdência e seguridade, a qual é grandemente impactada. Tais impactos estão diretamente ligados à necessidade de provisionamento de recursos necessários ao pagamento de rendas de sobrevivência – aposentadorias e pensões - por tempo superior àquele determinado através de tábuas de mortalidade estáticas, baseadas em experiências de outros países, podendo não refletir a mortalidade atual e futura da população brasileira, especialmente aquela parcela da população que recebe renda proveniente da Previdência Complementar Fechada – Fundos de Pensão.

Assim, dada a evolução contínua dos níveis de mortalidade da população – que alcançam patamares cada vez menores – a escolha das tábuas de mortalidade que melhor reflitam tais níveis, se torna uma tarefa de extrema relevância por parte das entidades de previdência e seguridade, tendo em vista que é um instrumento indispensável nos cálculos de rendas de sobrevivência, contribuições, prêmios, bem como o dimensionamento das Provisões Matemáticas.

As medidas que vêm sendo tomadas por parte dos sistemas de previdência e seguridade, com o objetivo de mensurar os níveis de mortalidade de seus participantes e avaliar a adequabilidade das tábuas de mortalidade adotadas, é a realização de testes de

aderência, os quais se baseiam nas populações expostas ao risco e informações inerentes a elas, tais como seu tamanho e as idades das pessoas que as compõem. Tais testes são realizados no mínimo uma vez ao ano e geralmente consideram aquela parcela da população que já esteja em gozo de benefício, isto é, os participantes assistidos. Além desta ferramenta de mensuração, também existem determinações de órgãos reguladores destes sistemas num sentido prudencial, estabelecendo parâmetros mínimos a serem adotados quanto à mortalidade de seus participantes. O Conselho de Gestão de Previdência Complementar (CGPC), já emitiu duas Resoluções: a primeira é a Resolução CGPC N° 18 de 28 de Março de 2006, que estabelece a adoção da tábua biométrica AT-1983 (*Annuity Table 1983*) como parâmetro mínimo de mortalidade, cuja expectativa de vida (ao nascer) é de aproximadamente 79 anos. Já a Resolução CGPC N° 26 de 29 de Setembro de 2008, que é a segunda, estabelece como tal parâmetro a adoção da tábua biométrica AT-2000 (*Annuity Table 2000*), desde que observadas as recomendações feitas na própria Resolução. Para esta última tábua, a expectativa de vida ao nascer é de pouco mais de 81 anos de idade.

Apesar das medidas adotadas pelos sistemas de previdência e seguridade, bem como a atitude prudencial do CGPC em estabelecer as tábuas já citadas como parâmetros mínimos de mortalidade para as Entidades de Previdência Complementar Fechadas (EFPC's), a qual também visa preservar o equilíbrio e solvência dos planos das EFPC's; surge a questão de quão fidedignamente tais tábuas estariam representando a real mortalidade dos participantes destas entidades, uma vez que se tratam de experiências internacionais. Uma tentativa de responder a tal pergunta, é a realização de testes de aderência conforme já citado. Todavia, a validade dos resultados da aplicação desta ferramenta, está atrelada ao pressuposto estatístico de que a população em estudo deve ser suficientemente grande, pois só assim o "princípio da aderência" é atendido. Ocorre que, em várias EFPC's, o tamanho das populações sobre as quais se aplicam tais testes, são inferiores a 3000 participantes, impossibilitando a afirmação de que o referido pressuposto tenha sido atendido. Assim, por se tratar de populações pequenas, as conclusões de tais testes acabam se fundamentando no "princípio da tendência", que como já é sabido e comprovado em estudos, é de queda nos níveis de mortalidade.

Neste sentido, o primeiro objetivo deste trabalho é a construção de uma tábua de mortalidade para idades específicas utilizando o procedimento de graduação de taxas de mortalidade, baseada nas experiências das EFPC's (considerando uma população de tamanho mínimo razoável), de forma a obter um comparativo dos níveis de mortalidade da população em estudo com aqueles provenientes das tábuas estabelecidas como parâmetros mínimos e outras que eventualmente são adotadas nas Entidades. Já o segundo objetivo, consiste em

obter fatores de melhoria – ou escalas de projeções – aplicando-os sobre as probabilidades de morte da tábua construída, tornando-a geracional. Assim, busca-se obter os comportamentos dos níveis futuros de mortalidade da população em estudo, de forma a possibilitar uma visão antecipada dos compromissos das entidades com seus participantes no futuro, mantendo a solvência e equilíbrio de seus planos.

A metodologia utilizada na construção da tábua consiste numa abordagem de inferência bayesiana em conjunto com a técnica de simulação estocástica de Monte Carlo via Cadeias de Markov (MCMC), por meio do Amostrador de Gibbs (*Gibbs Sampler*). Através de tais técnicas, foram implementados modelos paramétricos e não paramétricos de Graduação de Taxas Brutas de Mortalidade. Para cada modelo implementado, avaliou-se o melhor ajuste aos dados observados e melhor previsão dos dados replicados, obtendo-se a tábua de mortalidade resultante.

Fatores de melhoria – ou escalas de projeções – para tornar a tábua construída na forma geracional, foram obtidos mediante aplicação do modelo de Lee-Carter, também implementado estocasticamente utilizando *bootstrap* paramétrico num processo ARIMA(p,d,q), o qual se ajusta à componente  $k_t$  do modelo, a qual corresponde à uma série temporal que descreve a mudança na mortalidade ao longo do tempo.

Os dados disponibilizados para o estudo são provenientes de seis EFPC's do sul e sudeste do país, a saber, Fundação Copel de Previdência e Assistência Social, Fundação CEEE, FUNCESP, Fundação Itaipu BR de Previdência e Assistência Social (FIBRA), Fundação Celesc de Seguridade Social (CELOS) e Fundação Eletrosul de Previdência e Assistência Social (ELOS), tendo todas em comum o seguimento de energia elétrica. Para cada Entidade, os dados referem-se aos participantes assistidos (aposentados) no período compreendido entre os anos 2000 e 2008, considerando o mês de dezembro de cada ano.

Tanto na implementação dos modelos de Graduação Bayesiana como na simulação do processo ARIMA(p,d,q) por *bootstrap* paramétrico, foram utilizados os programas estatísticos R (R Development Core Team, 2009), JAGS (PLUMMER, 2009) e WinBUGS (SPIEGELHALTER; THOMAS; BEST; LUNN, 2003).

Os pacotes JAGS e WinBUGS, são versões da linguagem BUGS (*Bayesian Inference Using Gibbs Sampler*), por meio dos quais foram implementadas simulações estocásticas por MCMC através do Amostrador de Gibbs. O JAGS é desenvolvido para o sistema operacional Linux, podendo também ser executado sob o Windows, assim como o WinBUGS. Ambos foram executados a partir de programas escritos em linguagem R que se utilizam de interfaces disponibilizadas por pacotes adicionais do R, sendo que para o

WinBUGS em especial, é necessário a instalação do pacote R2WinBUGS (STURTZ; LIGGES; GELMAN, 2005). As análises e diagnósticos de convergência das cadeias foram realizados através do pacote coda (PLUMMER; BEST; COWLES; VINES, 2009), interno ao R. Os softwares são todos gratuitos de código aberto, estando o R disponível em (<http://cran.r-project.org/bin/windows/base/>), o JAGS em (<http://www-ice.iarc.fr/~martyn/software/jags/>) e o WinBUGS em (<http://www.mrc-bsu.cam.ac.uk/bugs/>).

Finalmente, a estruturação deste trabalho está organizado da seguinte forma: no capítulo 2 são abordadas as metodologias empregadas tanto na graduação como na projeção das taxas de mortalidade, bem como alguns conceitos básicos de atuária, os quais são essenciais. No capítulo 3 é realizada uma análise dos dados e são apresentadas as definições e aplicações dos modelos de Graduação Bayesiana, bem como a construção da tábua bayesiana de mortalidade juntamente com algumas considerações sobre os resultados alcançados. Já no capítulo 4, é verificada a adequabilidade da aplicação do modelo de Lee-Carter aos dados da população de expostos ao risco, visando tornar a tábua de mortalidade bayesiana na forma geracional, discutindo os resultados decorrentes da aplicação do modelo e as restrições de sua aplicação aos dados e, por fim, o capítulo 5 apresenta as conclusões e sugestões de extensão deste trabalho.

Todos os códigos de comandos construídos para implementação das análises, se encontram disponíveis no Apêndice.



## 2 METODOLOGIAS

Este trabalho está estruturado sobre o paradigma de inferência bayesiana para a graduação de taxas de mortalidade e a técnica de projeção de Lee-Carter (LEE; CARTER, 1992), ambos implementados através de métodos computacionalmente intensivos, especificamente simulação de Monte Carlo via Cadeias de Markov (MCMC) e *bootstrap* paramétrico, respectivamente.

No que diz respeito à Graduação de Taxas Brutas de Mortalidade, os modelos propostos são de estrutura hierárquica (NEVES, 2005), implementados por MCMC. Os modelos são construídos com base na função de verossimilhança e especificação das *prioris* dos parâmetros de interesse. Inferências são baseadas nas distribuições a *posteriori* dos mesmos. Na metodologia de projeção de taxas de mortalidade, o modelo de Lee-Carter, mescla um modelo demográfico com séries temporais, onde uma de suas componentes é ajustada por um processo ARIMA(p,d,q), o qual foi implementado de forma estocástica através de *bootstrap* paramétrico. Este capítulo revisa tais conceitos e técnicas, incluindo uma introdução a conceitos básicos de Atuária, os quais são fundamentais tanto na graduação como na projeção de taxas de mortalidade.

### 2.1 CONCEITOS BÁSICOS DE MATEMÁTICA ATUARIAL

Tanto na graduação das taxas brutas de mortalidade como na projeção das taxas graduadas, são utilizados conceitos de Atuária quando do desenvolvimento destas etapas do trabalho. As notações utilizadas são universalmente aceitas, buscando atender as convenções determinadas pelo *International Actuarial Association's Permanent Committee on Notation*.

Pode-se dizer que a probabilidade de uma pessoa com determinada idade  $x$ , morrer entre esta idade e a idade  $x+n$ , é o conceito a partir do qual se desenvolvem os demais. Tal probabilidade é dada na forma condicional, representada pela seguinte expressão:

$$P(X \leq x+n | X \geq x) = \frac{P(x \leq X < x+n)}{P(X \geq x)} = \frac{l_x - l_{x+n}}{l_x} = \frac{{}_n d_x}{l_x} = {}_n q_x \quad (2.1)$$

em que  $X$  é uma variável aleatória contínua que representa a idade de morte de um recém nascido. Na notação atuarial,  $l_x$  e  $l_{x+n}$  são respectivamente os números esperados de

sobreviventes às idades  $x$  e  $x+n$ , dentre o total de expostos ao risco no início da coorte hipotética ( $l_0$  recém nascidos), enquanto que  ${}_n d_x$  representa o número esperado de mortes entre as idades  $x$  e  $x+n$ . A probabilidade representada por  ${}_n q_x$ , é a proporção entre o número de mortes ocorridas entre as idades  $x$  e  $x+n$  e o total de vivos na idade  $x$ . O índice  $n$  geralmente é dado em anos completos, podendo também ser uma fração de ano.

Do complementar de (2.1), obtém-se a probabilidade representada por  ${}_n p_x$  de uma pessoa com idade  $x$  ultrapassar a idade  $x+n$ , isto é, sobreviver à idade  $x+n$ , dada pela seguinte expressão:

$$1 - P(X \leq x+n | X \geq x) = \frac{P(X > x+n)}{P(X \geq x)} = 1 - {}_n q_x = \frac{l_{x+n}}{l_x} = {}_n p_x \quad (2.2)$$

Nas expressões acima, quando  $n$  é igual a 1, usualmente omite-se o subscrito, isto é,  ${}_1 q_x$  e  ${}_1 p_x$  são representados simplesmente por  $q_x$  e  $p_x$ . Também, da definição da variável aleatória contínua  $X$ , e sendo  $F(x)$  sua função distribuição, tem-se:

$$F(x) = P(X \leq x), \quad x \geq 0$$

da qual podemos definir função de sobrevivência, denotada por  $S(x)$ , como sendo

$$S(x) = 1 - F(x) = P(X > x), \quad x \geq 0$$

Assim, por convenção, tem-se que  $F(0)=0$ , logo  $S(0)=1$ . Neste ponto, podemos introduzir outro conceito fundamental tanto para graduação como para projeção de taxas de mortalidade, denominado “força de mortalidade”, denotada por  $\mu_x$ . A aplicação deste conceito, não se dá somente em análise de sobrevivência, onde  $\mu_x$  é chamada variação instantânea da intensidade de morte, ou ainda, risco instantâneo de morte; mas também se aplica em teoria da confiabilidade, na qual é chamada taxa de falha. No contexto atuarial, a variação instantânea da intensidade de morte pode ser entendida como a probabilidade de uma pessoa de idade  $x$  – representada por  $(x)$  na notação atuarial – morrer dentro de um período  $t$  relativamente curto, que geralmente é menor que um ano. Neste caso,  $(x)$  morreria antes de atingir a idade  $x+t$ , com  $t$  sendo uma fração de ano. Logo, considerando que  $t \rightarrow 0$ , tal probabilidade ( $\mu_x$ ) pode ser representada em termos de  $F(x)$  e  $S(x)$ , da seguinte forma:

$$\begin{aligned}
\mu_x &= P(X < x+t | X \geq x) = \frac{P(X < x+t) - P(X < x)}{P(X \geq x)} \\
&= \frac{F(x+t) - F(x)}{1 - F(x)} \\
&= \frac{S(x) - S(x+t)}{S(x)} \\
&= -\frac{S'(x)}{S(x)}
\end{aligned} \tag{2.3}$$

Ainda, usando a notação de limite, a representação de  $\mu_x$  pode tomar a seguinte forma:

$$\mu_x = \lim_{z \rightarrow x^+} \frac{F(z) - F(x)}{1 - F(x)} = -\frac{S'(x)}{S(x)} \tag{2.4}$$

em que  $z - x = t \rightarrow 0$ . Também, uma vez definida  $\mu_x$ , pode-se expressar  ${}_tq_x$  e  ${}_tp_x$  (com  $t$  suficientemente pequeno) em função da mesma, conforme segue:

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

e

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

Especificamente no contexto de graduação - a qual se dará em função da força de mortalidade - como o objetivo é obter probabilidades anuais de morte em cada idade  $x$  ( $q_x$ ), trabalha-se com a seguinte relação (BOWERS, 1997):

$$q_x = 1 - \exp(-\mu_x) \tag{2.5}$$

logo, tem-se que

$$\mu_x = -\ln(1 - q_x) \tag{2.6}$$

De acordo com a simplificação contida em Neves (2005), considera-se  $\mu_x$  constante em cada faixa ou intervalo de um ano de idade, a partir de uma certa idade  $x$ , isto é, entre  $x$  e  $x+1$ ,  $x+1$  e  $x+2$ , etc. Assim,  $\mu_i = \theta_i$  e assume-se que  $\theta_i \leq \theta_{i+1}$ , em que  $i=1, \dots, k$  representam as classes de idade, sendo  $k$  a última classe. Substituindo em (2.5), tem-se:

$$\begin{aligned} \mu_y &= \theta_i, \quad x+i-1 \leq y \leq x+i \text{ e} \\ q_{x+i-1} &= 1 - \exp(-\theta_i), \text{ com classes de idade } i = 1, \dots, k \end{aligned} \quad (2.7)$$

Portanto, nos modelos que serão empregados na graduação, as probabilidades de morte obedecem à relação (2.7). Também, a equação (2.6) será empregada na obtenção de forças de mortalidade a partir das probabilidades de morte  $q_x$  obtidas de tábuas de mortalidade comumente adotadas nas EFPC's.

Uma vez construída a tábua de mortalidade através do modelo de graduação escolhido, será avaliado o impacto nas Reservas (Provisões) Matemáticas, decorrente da aplicação das taxas de mortalidade graduadas no dimensionamento de tais reservas. Tal dimensionamento se dará através da multiplicação do valor da renda mensal (anualizada) de um assistido (aposentado) pelo valor de uma anuidade mensal, a qual é calculada com base na probabilidade de morte ( $q_x$ ) do referido assistido, na respectiva idade  $x$ . Esta anuidade é definida como o valor presente atuarial de uma renda aleatória vitalícia, de valor unitário, postecipada e fracionada em  $m$  períodos, paga a uma pessoa de idade  $x$ , sendo dada pela seguinte expressão:

$$a_x^{(m)} = \frac{N_x}{D_x} - \left( \frac{m+1}{2m} \right) = \frac{\sum_{k=0}^{\overline{m}} D_{x+k}}{D_x} - \left( \frac{m+1}{2m} \right)$$

em que  $D_x = l_x(1+j)^{-x}$  e  $N_x = \sum_{k=0}^{\overline{m}} D_{x+k}$  são chamados comutações,  $l_x$  já foi definido anteriormente,  $x$  representa a idade,  $j$  é a taxa de desconto atuarial e  $m$  representa o número de períodos em que é fracionado o ano para efeito dos pagamentos da renda vitalícia. A taxa de desconto atuarial  $j$  merece destaque, estudos na área econômica comprovam que a tendência é de queda na taxa de juro, de forma que já existem algumas EFPC's adotando taxas de desconto atuarial de 5% ou 5,5% ao ano, por conservadorismo. Todavia, a maioria das entidades de previdência e seguridade ainda adota esta taxa como sendo 6% ao ano, razão pela qual também consideraremos este valor neste estudo.

Neste trabalho, os expostos ao risco recebem rendas mensais, logo o número de fracionamentos do ano é igual a 12, isto é,  $m=12$ . Portanto, temos que:

$$a_x^{(12)} = \frac{N_x}{D_x} - \frac{13}{24}$$

A referida anuidade também será utilizada na realização do teste de aderência para avaliação da adequabilidade das tábuas AT-1949, AT-1983 e AT-2000 à mortalidade da população exposta ao risco.

## 2.2 GRADUAÇÃO DE TAXAS DE MORTALIDADE

Quando da construção de uma tábua de mortalidade com base na experiência de uma certa população exposta ao risco, primeiramente calcula-se as taxas brutas de mortalidade em todas as idades para esta população, através da seguinte equação:

$$v_x = \frac{l_x^o - l_{x+1}^o}{e_x} = \frac{d_x^o}{e_x} \quad (2.8)$$

em que  $x$  é a idade de exposição ao risco,  $e_x$  é a quantidade central de expostos ao risco observados nesta idade,  $d_x^o$  representa o número de mortes observado entre as idades  $x$  e  $x+1$ , sendo  $l_x^o$  e  $l_{x+1}^o$  os números de expostos ao risco observados nas idades  $x$  e  $x+1$ , respectivamente. A quantidade central de expostos ao risco é dada por:

$$e_x = \frac{l_x^o + l_{x+1}^o}{2}$$

O comportamento das taxas brutas de mortalidade calculadas, geralmente é um tanto errático, indicando ausência de suavidade e monotonicidade crescente em relação à idade, dado que este é o comportamento esperado da mortalidade humana a partir de certa idade. Para induzir este comportamento, utiliza-se um processo de suavização das taxas brutas de mortalidade que, além de evitar o crescimento brusco das taxas de mortalidade, visa manter a monotonicidade desejada das mesmas, bem como contornar o problema da falta de

informação nas idades em que não se observou expostos ao risco. Este processo é chamado *graduação de taxas brutas de mortalidade* e, através dele, são obtidas as probabilidades de morte  $q_x$  da população em cada idade, as quais irão compor a tábua de mortalidade, que é o instrumento indispensável em cálculos e avaliações atuariais.

Os modelos de graduação de taxas de mortalidade são classificados em paramétricos e não paramétricos, isto é, globais e locais. No primeiro caso, pressupõe-se que a mortalidade do grupo de expostos ao risco obedece alguma lei de mortalidade, na qual se baseia alguma função de sobrevivência válida para todas as idades. Daí o fato de serem chamados modelos paramétricos ou globais. Já os modelos locais têm como fator principal o peso dos dados, não obedecendo a uma lei de mortalidade específica.

Também, a quantidade de expostos ao risco em determinada idade influencia o ajuste do modelo. Se esta quantidade for bastante reduzida para uma idade específica, isso faz com que se aumente a incerteza sobre a taxa de mortalidade graduada naquela idade.

### 2.2.1 MODELOS BAYESIANOS DE GRADUAÇÃO

Os modelos de graduação utilizados neste trabalho - tanto locais como globais - possuem a estrutura hierárquica bayesiana, os quais são propostos em Neves (2005), cujo objetivo é graduar as taxas brutas de mortalidade em função da força de mortalidade, as quais são obtidas a partir das tábuas biométricas usualmente adotadas nas EFPC's, através da aplicação da relação (2.6). Os modelos locais de graduação bayesiana são baseados nos modelos de Broffit (1998) e Carlin (1992) conforme descrito em Neves (2005), ao passo que os modelos globais são baseados nas leis de mortalidade de Gompertz e Makeham, adaptados para um enfoque bayesiano (NEVES, 2005).

A abordagem bayesiana do processo de graduação consiste na inferência sobre os parâmetros desconhecidos, sendo agregados aos dados alguns conhecimentos iniciais através da atribuição de distribuições *a priori* aos parâmetros de interesse, que neste caso são as forças de mortalidade. Desta forma, todos os parâmetros são tratados como quantidades aleatórias e as inferências são feitas sobre a distribuição *a posteriori* dos parâmetros de interesse, a partir da qual são calculadas as probabilidades de morte  $q_x$  de acordo com (2.7). O processo de inferência neste caso é implementado por um algoritmo de MCMC que fornece amostras da distribuição *a posteriori* dos parâmetros. O algoritmo MCMC consiste em combinar a verossimilhança Poisson para as mortes observadas  $d_i$  com as *prioris* para  $\theta_i$ .

Com isto a cada  $r$ -ésima iteração são geradas amostras  $d_i^r$  da preditiva do número de mortes. A partir destes valores, as forças de mortalidade futuras  $\theta_i^r = d_i^r / e_i$  e as probabilidades de morte futuras  $q_{x+i-1}^r = 1 - \exp(-\theta_i^r)$  são calculadas para as idades  $x, x+1, \dots, x+k-1$ , com  $i=1, \dots, k$  representando as classes de idades. Estas probabilidades são o objetivo final do processo de graduação. Também, para todos os modelos de graduação, é assumido que  $d_i | \theta_i$  segue uma distribuição de Poisson( $e_i \theta_i$ ), com  $e_i$  sendo a quantidade central de expostos ao risco para a classe de idade  $i$ .

Finalmente, os estimadores bayesianos mais comuns para os parâmetros de interesse, são a média, moda e mediana da distribuição *a posteriori* destes parâmetros. Neste trabalho, o estimador das probabilidades de morte futuras será a esperança da distribuição preditiva que, avaliada no contexto de inferência por MCMC, é simplesmente a média das amostras da *posteriori*.

Desta forma, as inferências são inteiramente baseadas nas distribuições *a posteriori*, constituindo a essência da inferência bayesiana, que em resumo, consiste de quatro aspectos fundamentais:

- **Informação a *Priori*:** Todos os problemas são únicos e possuem seus próprios contextos. Tais contextos permitem deduzir informação a priori, sendo a formulação e explicação de tal informação, o que estabelece um aspecto particular da inferência bayesiana.
- **Probabilidade Subjetiva:** A estatística clássica depende da definição frequentista de probabilidades, a qual está alicerçada na realização de um número suficientemente grande de vezes de um experimento. Por outro lado, a inferência bayesiana formaliza explicitamente a noção de que todas as probabilidades subjetivas dependem da crença individual e conhecimento disponível. Portanto, uma análise bayesiana é personalizada de acordo com as especificações da crença a “*priori*” de cada indivíduo, a respeito do parâmetro de interesse  $\theta$ . Assim, a inferência é realizada com base na distribuição *a posteriori* de  $\theta$  a partir da especificação de uma distribuição *a priori* de  $\theta$ .
- **Consistência Própria:** O tratamento do parâmetro  $\theta$  como sendo aleatório, expressando assim a incerteza sobre este, implica que todo o desenvolvimento da

inferência bayesiana surge apenas da teoria de probabilidade. Isto tem muitas vantagens, e significa que todas as questões a respeito da inferência, podem ser direcionadas como afirmações probabilísticas sobre  $\theta$ , que então se deduz diretamente da distribuição a *posteriori*.

- **Ausência de Procedimento *ad hoc***: No contexto de inferência clássica não se pode fazer afirmações probabilísticas sobre o parâmetro  $\theta$ . Vários critérios são desenvolvidos para julgar se um particular estimador é em algum sentido “bom”. Isto tem levado à proliferação de procedimentos, freqüentemente conflitantes. A inferência bayesiana descarta esta tendência de criar critérios *ad hoc*, julgando e comparando estimadores com base na distribuição a *posteriori*, para expressar apenas em termos probabilísticos, toda a inferência sobre o parâmetro  $\theta$  desconhecido.

Na implementação de inferência bayesiana, a obtenção de *posteriors* é facilitada quando estas possuem distribuição na mesma família da *priori*, sendo necessária apenas alguma atualização dos parâmetros da distribuição a *priori* com base na informação contida nos dados. Tal fato caracteriza a “conjugação”. A Tabela 2.1 contém algumas famílias de distribuições a *priori* conjugadas.

Tabela 2.1 - Famílias padrão de distribuições a *priori* conjugadas

Verossimilhança	Priori	Posteriori
$X \sim \text{Binomial}(n, \theta)$	$\text{Beta}(p, q)$	$\text{Beta}(p + x, q + n - x)$
$X_1, \dots, X_n \sim \text{Poisson}(\theta)$	$\text{Gama}(p, q)$	$\text{Gama}(p + \sum_{i=1}^n x_i, q + n)$
$X_1, \dots, X_n \sim N(\theta; \tau^{-1}), \tau \text{ conhecido}$	$N(b, c^{-1})$	$N\left(\frac{cb + nr\bar{x}}{c + n\tau}, \frac{1}{c + n\tau}\right)$
$X_1, \dots, X_n \sim \text{Gama}(k, \theta), k \text{ conhecido}$	$\text{Gama}(p, q)$	$\text{Gama}(p + nk, q + \sum_{i=1}^n x_i)$
$X_1, \dots, X_n \sim \text{Ge}(\theta)$	$\text{Beta}(p, q)$	$\text{Beta}(p + n, q + \sum_{i=1}^n x_i - n)$
$X_1, \dots, X_n \sim \text{BinomialNegativa}(r, \theta)$	$\text{Beta}(p, q)$	$\text{Beta}(p + r, q + x - r)$

## 2.2.2 IMPLEMENTAÇÃO DOS MODELOS DE GRADUAÇÃO BAYESIANA



Os modelos de graduação bayesiana serão implementados com o uso de métodos computacionalmente intensivos, utilizando simulação estocástica de MCMC por meio do Amostrador de Gibbs. Através dos métodos computacionalmente intensivos, que neste caso se classificam em computação para inferência (COLES, 2007), foi possível a avaliação de integrais complexas decorrentes dos modelos em questão, representadas pela constante normalizadora na aplicação do Teorema de Bayes. Essa complexidade torna inviável a realização de inferência exata para a obtenção das *posteriors* das taxas graduadas.

O método de simulação de Monte Carlo via Cadeias de Markov (MCMC), é utilizado aqui para gerar amostras a partir de uma distribuição fácil de ser amostrada, devido à impossibilidade de se amostrar diretamente da distribuição de interesse. No caso do MCMC via Amostrador de Gibbs, amostra-se das distribuições condicionais completas. O método é iterativo, baseado em cadeias de Markov, de forma que as observações geradas não são independentes. Todavia, isto não causa problemas, devido ao grande número de iterações realizadas, bem como o intervalo entre as observações retidas no processo de simulação (*thinning*) que, dependendo de sua frequência, diminui a autocorrelação entre as observações.

Uma vez obtidas as amostras da distribuição de interesse, se torna possível calcular estimativas amostrais de características desta distribuição. No caso das probabilidades de morte estimadas  $\hat{q}_x$ , estas são dadas pelas médias dos valores  $q_x^r$  obtidos na simulação.

As cadeias de Markov são geradas através do algoritmo Metropolis-Hastings (HASTINGS; 1970), por um longo período, de forma que sejam:

- homogêneas: as probabilidades de transição de um estado para o outro são constantes;
- irredutíveis: cada estado pode ser atingido a partir de qualquer outro em um número finito de iterações;

Um caso especial do algoritmo Metropolis-Hastings (MH) é o Amostrador de Gibbs (GEMAN & GEMAN, 1984), o qual já vem sendo utilizado há bastante tempo, implementado em *softwares* como o WinBUGS e JAGS. Nestes programas, caso seja possível usar exclusivamente um algoritmo de Amostrador de Gibbs, não há necessidade de calibração através dos parâmetros da *proposal*, uma vez que o uso de condicionais completas elimina o passo de rejeição/aceitação característico do MH. Cada novo valor gerado através do Amostrador de Gibbs é sempre aceito na cadeia e as transições de um estado para outro são feitas de acordo com as distribuições condicionais completas, que neste caso, correspondem

às distribuições das probabilidades de morte futuras  $q_{x+i-1}^r$ , com  $i=1, \dots, k$ , para as idades  $x, x+1, \dots, x+k-1$ .

A convergência das cadeias geradas por MCMC, não é automática. Primeiramente é necessário certo número de realizações da cadeia até obter a convergência para a distribuição estacionária. Essas primeiras realizações são descartadas posteriormente, servindo apenas para estabilizar a cadeia, sendo este período conhecido como *burn-in* (período de aquecimento), e seu tamanho (número de realizações) dependerá da calibração do algoritmo. Alguns programas podem fazer a calibração internamente e, o número de realizações a título de *burn-in*, geralmente é adotado como sendo a metade do total de iterações. Entretanto, não há garantia de convergência através desses critérios automáticos e a calibração adequada deve ser investigada pelo usuário. Para isso, é usual a realização de testes de diagnósticos de convergência das cadeias, como por exemplo os implementados no pacote coda do programa R.

### 2.3 O MODELO DE LEE-CARTER

O modelo de Lee-Carter (LEE; CARTER, 1992), é o método de projeção de taxas de mortalidade adotado neste trabalho. Busca-se aqui obter os fatores de melhoria (*improvement*) para projeções das taxas de mortalidade graduadas futuras, com base na experiência das EFPC's participantes deste estudo. Trata-se de um método extrapolativo, isto é, baseado na projeção de tendências históricas e padrões etários da mortalidade para o futuro, incluindo algum elemento de julgamento subjetivo, como por exemplo, a escolha do período sobre o qual as tendências serão determinadas.

Este método foi apresentado por Ronald Lee e Lawrence Carter no início da década de 90, aplicado inicialmente em dados de mortalidade da população americana no período de 1900 a 1989, com um horizonte de previsão de 76 anos. Desde então, adquiriu certa popularidade, tendo sido aplicado com sucesso não só nos EUA, mas também em outros países como: Canadá, Chile, Japão, França, Alemanha, dentre outros (CMIB, 2004). No Brasil, também já foi aplicado em dados da população (FIGOLI, 1998), e recentemente, utilizado em um estudo para obtenção de escalas de projeção (fatores de *improvement*) com o intuito de construir previsões de mortalidade com base na experiência de mortalidade da população da Inglaterra e País de Gales (SANTOS, 2007). Girosi e King (2007) apresentam uma revisão crítica da metodologia proposta por Lee e Carter (1992), explorando

propriedades do modelo e discutindo sua estimação, adequação e implementação em diversos contextos.

O método é uma combinação de um modelo demográfico bilinear nas variáveis  $x$  (idade) e  $t$  (ano calendário) com um modelo de séries temporais, permitindo a obtenção de intervalos probabilísticos para as respectivas previsões. O modelo pode ser estimado a partir de uma matriz de taxas específicas de mortalidade por idade, para diversos períodos passados, obtendo-se um índice temporal do nível geral da mortalidade. Embora simples e parcimonioso, é altamente estruturado, tipicamente capturando quase toda a variação nos dados (CMIB, 2004). Algumas de suas vantagens são as seguintes:

- permite que as taxas de mortalidade decresçam de forma exponencial ao longo tempo, não sendo necessário estabelecer um limite superior arbitrário ou racionalizar de alguma forma a desaceleração dos ganhos na esperança de vida;
- é possível se obter de forma indireta a mortalidade para períodos nos quais não se dispõe de taxas de mortalidade específicas, mediante extrapolação.

O método também apresenta limitações, pois não incorpora:

- informação sobre avanços na medicina;
- alteração do contexto sócio econômico;
- alteração de estilo de vida;
- informação sobre o aparecimento de novas doenças;

Mesmo com as limitações mencionadas, o método tem sido aplicado em estudos de mortalidade por causa específica em muitos países e em vários períodos de tempo, extrapolando seu contexto inicial de aplicação. Contudo, várias destas aplicações não têm sentido (GIROSI; KING, 2007).

O método consiste em decompor a mortalidade por idade em duas componentes, um índice de mortalidade variando no tempo,  $k_t$ , e um conjunto de coeficientes para as idades.

O modelo tem a forma:

$$\log m_{x,t} = a_x + b_x k_t + \varepsilon_{x,t}, \text{ com } x = 1, \dots, X \text{ e } t = 1, \dots, T \quad (2.9)$$

em que  $m_{x,t}$  é a taxa de mortalidade na idade  $x$  e no ano calendário  $t$ ;  $\varepsilon_{x,t}$  é um erro aleatório com média zero e variância  $\sigma_\varepsilon^2$ ; os coeficientes  $a_x$  descrevem a forma geral do perfil de mortalidade por idade; os coeficientes  $b_x$  descrevem o padrão de desvios do perfil da idade conforme o parâmetro  $k_t$  varia. Se o coeficiente  $b_x$  é particularmente alto para algumas idades  $x$ , isto significa que as taxas de mortalidade mudam mais rapidamente nessas idades. Se os  $b_x$  são iguais, então as taxas de mortalidade mudam à mesma velocidade em todas as idades.

O parâmetro  $k_t$  corresponde a uma série temporal que descreve a mudança na mortalidade ao longo do tempo. Se  $k_t$  diminui, então as taxas de mortalidade diminuem, e se aumenta, as taxas de mortalidade aumentam. Os coeficientes  $b_x$  determinam como esta mudança global na mortalidade afeta as taxas na idade em questão, isto é:

$$\frac{\partial \log(m_{x,t})}{\partial t} = \frac{\partial k_t}{\partial t} b_x$$

Se  $k_t$  decresce linearmente, então  $m_{x,t}$  decresce exponencialmente em cada idade, a uma taxa que depende de  $b_x$ , a menos que  $b_x$  seja negativo, caso no qual  $m_{x,t}$  aumenta.

O modelo não especifica escolhas únicas para os parâmetros, pois  $b_x$  e  $k_t$  aparecem apenas através de seu produto  $b_x k_t$ . Frequentemente impõem-se as restrições  $\sum_x b_x = 1$  e  $\sum_t k_t = 0$ , para impor unicidade.

Lee e Carter também sugeriram o uso de um modelo de séries temporais para modelar  $k_t$ . As projeções baseadas em um modelo de Lee-Carter compartilham muitas das características de previsões em séries temporais. Uma vez ajustado o modelo,  $k_t$  é projetado de forma a dar as taxas de mortalidade para anos futuros. Uma escolha popular para a modelagem de  $k_t$  é o modelo ARIMA(0,1,0). O fato dessa componente ser modelada por processos da classe ARIMA (BOX; JENKINS, 1976), pressupõe que o histórico passado no qual o modelo se baseia para fazer extrapolações, seja suficientemente grande. Logo, as previsões podem estar comprometidas se o histórico dos dados não for relativamente longo. Booth *et al* (2006) comparam variações e extensões de previsão de mortalidade aplicando o

método em populações de 10 países e não encontram diferenças entre os métodos considerados. Estes métodos são: o modelo de Lee-Carter em sua forma clássica mais 4 variações do modelo, sendo a primeira segundo Lee-Miller, a segunda de acordo com Booth-Maindonald-Smith, a terceira segundo Hyndmann-Ullah e a quarta conforme De Jong-Tickle.

A seguir é descrito o método de ajuste do modelo adotado neste trabalho. Girosi e King (2007) apresentam e discutem alternativas de ajuste.

### 2.3.1 AJUSTE DO MODELO

Usualmente, o processo de ajuste do modelo demográfico é efetuado em duas etapas: estimação dos parâmetros  $a_x$ ,  $b_x$  e  $k_t$  e posterior reestimação de  $k_t$ , para que o número de óbitos ajustado coincida com o número de óbitos efetivamente verificado.

Uma solução exata de mínimos quadrados para os parâmetros  $a_x$ ,  $b_x$  e  $k_t$  é obtida através da aplicação do método da decomposição em valores singulares (*Singular Value Decomposition – SVD*). A SVD é aplicada à matriz dos logaritmos das taxas de mortalidade, após lhe ter sido subtraída a média temporal do logaritmo das taxas de mortalidade específicas por idade. Os primeiros vetores direito e esquerdo e o respectivo valor singular provenientes da aplicação do SVD à matriz, após a normalização, fornecem uma solução única para (2.9).

Uma vez geradas as estimativas para  $a_x$ ,  $b_x$  e  $k_t$  com base na matriz dos logaritmos das taxas de mortalidade e não na matriz de taxas de mortalidade; o número total de óbitos obtido pela aplicação das taxas de mortalidade estimadas à distribuição por idade da população, em geral, não coincide exatamente com o número total de óbitos observados. Lee e Carter (1992) introduziram uma nova etapa na estimação de  $k_t$ , ajustando-o de forma a que, para a estrutura etária da população e para os coeficientes  $a_x$  e  $b_x$  previamente estimados, produza exatamente o número total de óbitos observados para o ano em questão. Logo

$$D_t = \sum_x [(N_{x,t} \exp(a_x + k_t b_x))] \quad (2.10)$$

em que  $D_t$  é o número total de óbitos observado no ano  $t$  e  $N_{x,t}$  corresponde à população exposta ao risco da idade  $x$  no ano  $t$ .

### 2.3.2 AJUSTE DA COMPONENTE $k_t$ DO MODELO

Uma vez ajustado o modelo demográfico, o próximo passo é modelar  $k_t$ , como um processo estocástico de séries temporais. Para tanto, Lee e Carter (1992) utilizaram a metodologia Box & Jenkins (BOX; JENKINS, 1976), através de um modelo ARIMA de ordem (0,1,0), ou seja, um passeio aleatório com um termo constante,  $k_t = c + k_{t-1} + \gamma_t$ .

Com base nas estimativas do modelo selecionado, obtém-se a previsão do índice de nível de mortalidade,  $k_t$ , para um dado horizonte temporal. Conjuntamente com a previsão pontual, são obtidos intervalos de probabilidade para a previsão de  $k_t$ . Este é um dos aspectos mais importantes do método de Lee-Carter, uma vez que permite a construção de tais intervalos.

### 2.3.3 SIMULAÇÃO DA COMPONENTE $k_t$ E TAXAS BRUTAS DE MORTALIDADE

Neste trabalho serão investigados os resultados da aplicação do modelo de Lee-Carter ajustado aos dados da Inglaterra e País de Gales - considerando o período entre 1950 e 2003 - na projeção de taxas de mortalidade da população objeto deste estudo. O ajuste do modelo a esta experiência estrangeira, é devido à indisponibilidade de um histórico de dados suficientemente grande e que reflita a experiência das EFPC's. Logo, para esta avaliação, serão realizadas simulações da componente  $k_t$  do modelo de Lee-Carter ajustado, através de *bootstrap* paramétrico aplicado ao modelo ARIMA( $p,d,q$ ) que se ajusta a esta componente.

O bootstrap permite a obtenção da variabilidade esperada segundo o modelo implementado como descrito a seguir. Para cada valor de  $k_t$  simulado, será obtida uma sequência de taxas centrais de mortalidade  $m_{x,t}$ , uma para cada idade, conforme segue:

$$\begin{aligned} k_t^{(1)} &\rightarrow m_{x,t}^{(1)}, m_{x+1,t}^{(1)}, \dots, m_{x+j,t}^{(1)} \\ k_t^{(2)} &\rightarrow m_{x,t}^{(2)}, m_{x+1,t}^{(2)}, \dots, m_{x+j,t}^{(2)} \\ &\dots \\ k_t^{(r)} &\rightarrow m_{x,t}^{(r)}, m_{x+1,t}^{(r)}, \dots, m_{x+j,t}^{(r)} \end{aligned}$$

Logo, cada sequência de taxas centrais de mortalidade será gerada para as idade  $x, x+1, \dots, x+j$ , sendo  $k_t^{(1)}, k_t^{(2)}, \dots, k_t^{(r)}$  os valores simulados de  $k_t$ , com  $r$  sendo o total simulações e  $t$  um ano calendário específico. Desta forma, para cada ano calendário  $t$ , serão geradas  $r$  sequências de

taxas centrais de mortalidade  $m_{x,t}$ , sendo que das  $r$  sequências geradas, serão tomados os valores máximo e mínimo de cada taxa central gerada, obtendo-se os limites superior e inferior de um intervalo de confiança empírico para as taxas centrais de mortalidade.

Após a obtenção destes intervalos, serão comparadas as taxas centrais de mortalidade obtidas para a população em estudo, com aquelas geradas no processo de simulação. Caso as taxas centrais de mortalidade da população em estudo se situem entre os limites superior e inferior do intervalo de confiança - considerando um ano  $t$  específico - então o modelo de Lee-Carter ajustado a experiência inglesa será razoável na projeção de taxas de mortalidade da população em estudo.

Quanto ao número de simulações para geração dos intervalos de confiança empíricos, Marriott (1979) discutiu os aspectos técnicos relacionados ao número necessário de acordo com níveis de confiança pré-estabelecidos. Uma regra consagrada pela prática é a de se usar 19 simulações para 90% de confiança, 99 simulações para 95% de confiança e 499 simulações para 99% de confiança. Neste estudo serão usadas 99 simulações, visando obter intervalos de confiança empíricos de 95%.

### 3 GRADUAÇÃO DAS TAXAS BRUTAS DE MORTALIDADE

Os modelos de graduação utilizados neste trabalho são aqueles propostos em Neves (2005), os quais estão alicerçados sobre a inferência bayesiana e implementados de forma estocástica. Assim, este capítulo está dividido em duas partes: análise exploratória dos dados, e cálculo das taxas brutas de mortalidade, seguido das definições e aplicações dos modelos de graduação bayesiana, construção da tábua bayesiana de mortalidade e algumas considerações sobre o uso da metodologia no contexto dos dados tratados aqui.

Os dados utilizados são provenientes de seis EFPC's do setor elétrico: Fundação Copel de Previdência e Assistência Social (FC), Fundação CEEE, FUNCESP, Fundação Itaipu BR de Previdência e Assistência Social (FIBRA), Fundação Celesc de Seguridade Social (CELOS) e Fundação Eletrosul de Previdência e Assistência Social (ELOS), tendo todas em comum o seguimento de energia elétrica. Para cada Entidade, os dados referem-se aos participantes assistidos (aposentados) no período compreendido entre os anos 2000 e 2008, considerando o mês de dezembro de cada ano. Ainda, os participantes considerados foram aqueles que se aposentaram por tempo de serviço, idade, especial e outras aposentadorias não decorrentes de invalidez, de forma a obter as taxas de mortalidade geral da população.

A congregação dos dados das seis entidades visou obter uma população de tamanho razoável, buscando-se obter um estudo mais consistente sobre a mortalidade dos participantes das mesmas. Com este tamanho de população, é possível a graduação das taxas de mortalidade agregando ambos os sexos (unisex), devido a uma quantidade grande o bastante de expostos ao risco e mortes observadas. Contudo, quando separados por sexo, o número de expostos ao risco do sexo feminino é bastante reduzido, conforme mostra a Figura 3.3, não se observando mortes na maioria das idades, de forma que se torna impossível graduar as taxas de mortalidade para o sexo feminino.

Da mesma forma que em Neves (2005), trabalhou-se com a quantidade central de expostos ao risco  $e_x$  em cada idade e ano de estudo. Também, devido ao controle rigoroso da movimentação da população, o único motivo de exclusão de um participante dentre os expostos ao risco é a ocorrência do evento morte. O número de óbitos apurado em cada ano de estudo, se dá através da comparação entre o número de expostos no mês de dezembro de um determinado ano com o número de expostos em dezembro do ano subsequente. Desta forma, se uma pessoa entrar em aposentadoria num determinado ano, a partir do mês de



janeiro, vindo a falecer antes de dezembro do mesmo ano, e mesma será ignorada, não sendo considerada entre os expostos ao risco nem entre os óbitos ocorridos. Todavia, situações como esta são raras.

### 3.1 ANÁLISE DOS DADOS

A Tabela 3.1 resume os dados obtidos das EFPC's indicando a exposição, entradas e saídas anuais.

Tabela 3.1 - Movimentação de expostos ao risco no início, meio e fim de cada ano, agregando todas as EFPC's e não segregado por sexo.

Ano	Expostos início do ano	Novos entrados	Obitos ocorridos	Expostos fim do ano	Expostos meio do ano
2001	34627	922	386	35163	34895
2002	35163	477	383	35257	35210
2003	35257	2510	387	37380	36319
2004	37380	549	462	37467	37424
2005	37467	557	426	37598	37533
2006	37598	881	379	38100	37849
2007	38100	754	448	38406	38253
2008	38406	644	474	38576	38491

Observe-se que há um aumento substancial no número de novas entradas no ano de 2003. Este aumento é devido à inclusão dos dados da Fundação Eletrosul (ELOS) no estudo somente a partir de dezembro de 2003, elevando o número de expostos ao risco no início de 2004. O motivo foi a privatização de parte da companhia Eletrosul, a qual foi incorporada pela empresa belga Tractebel. Desta forma, os dados da ELOS utilizados aqui, se referem à parte da companhia que não foi privatizada, disponíveis somente a partir de 2003.

Durante o ano de 2006, também se observou um aumento significativo no número de novos entrados, embora não tão expressivo quanto aquele ocorrido em 2003. Este aumento é devido à concessão de novas aposentadorias pela FIBRA e Fundação Copel, que em 2006 atingiu patamares superiores aos anos anteriores. Anteriormente ao ano de 2006, eram concedidas em média 22,8 e 48 novas aposentadorias por ano pela FIBRA e Fundação Copel, respectivamente. Este aumento verificado em 2006 foi causado pela instituição de Planos de Demissões Incentivadas pelos patrocinadores destas entidades. Todavia, nos anos de 2007 e 2008, o número de novas aposentadorias para estas duas entidades, voltou a se comportar próximo aos patamares verificados anteriormente a 2006.

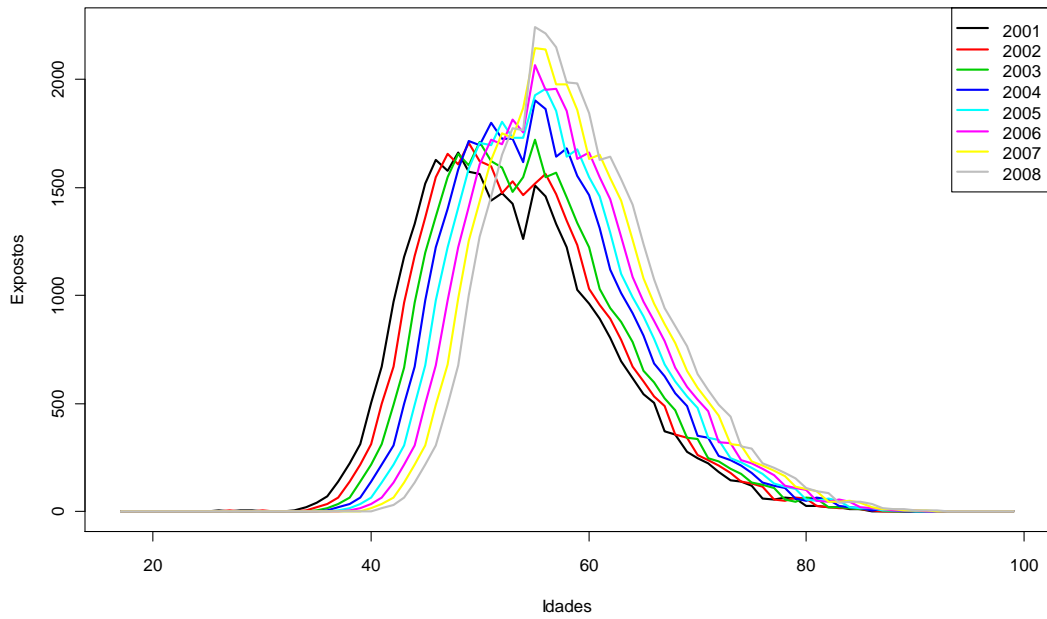


Figura 3.1 - Comportamento dos expostos ao risco de morte por idade e ano de estudo, para ambos os sexos, considerando todos os anos e entidades.

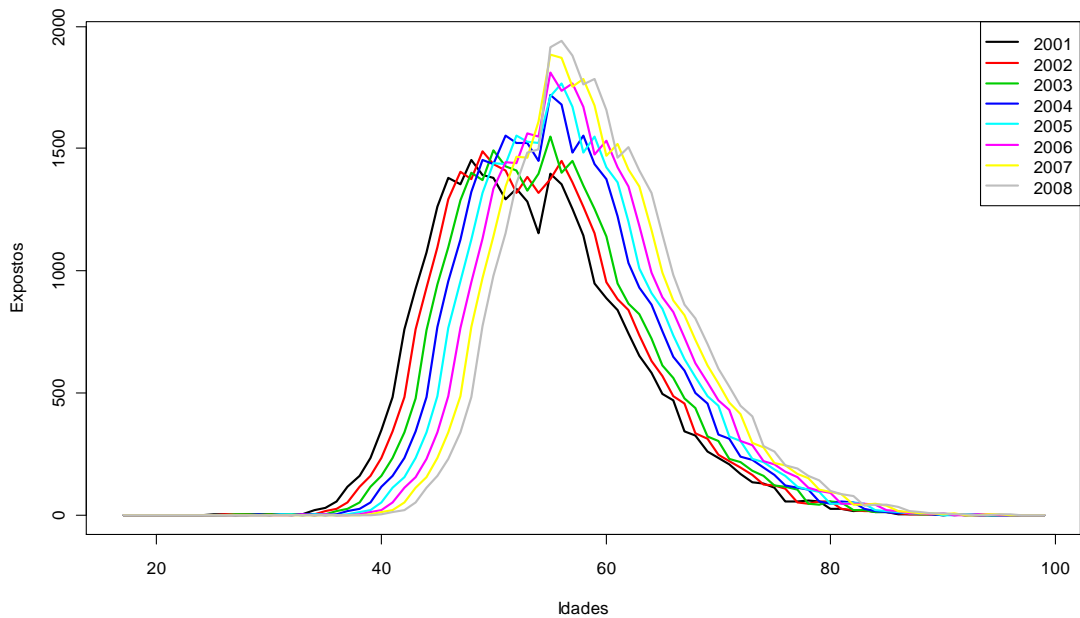


Figura 3.2: Comportamento dos expostos ao risco de morte por idade e ano de estudo, do sexo masculino, considerando todos os anos e entidades.

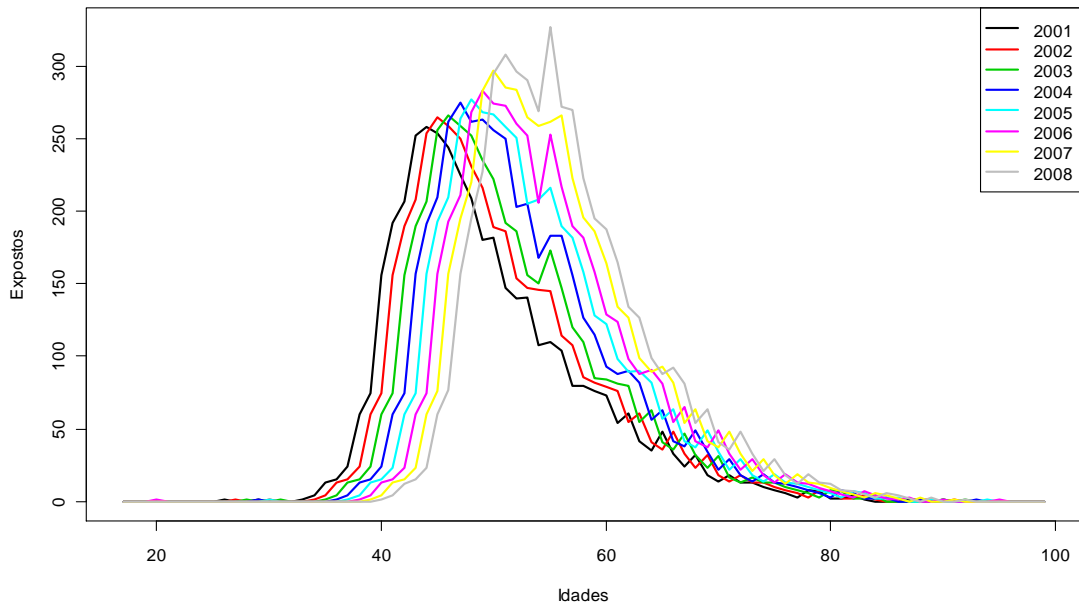


Figura 3.3: Comportamento dos expostos ao risco de morte por idade e ano de estudo, do sexo feminino, considerando todos os anos e entidades.

As Figuras 3.1, 3.2 e 3.3 apresentam o comportamento dos expostos ao risco segundo a idade, para ambos os sexos, somente para o sexo masculino e para o sexo feminino, respectivamente. Evidencia-se nas três figuras, a elevação do número de expostos ao risco nos anos de 2004 a 2008, se comparado aos anos de 2001 a 2003, devido à inclusão da ELOS no estudo no final do ano de 2003. Ainda, independentemente deste aumento no número de expostos, percebe-se que as idades modais de exposição ao risco geralmente são 55 anos - para unissex e sexo masculino - com exceção dos anos de 2001 e 2002, as quais são 48, 49 anos, respectivamente. Já para o sexo feminino, a idade modal de exposição se torna mais elevada à medida que se avança nos anos. A Tabela 3.2 apresenta as idades modais com seus respectivos números de expostos.

Tabela 3.2 - Quantidade de expostos ao risco em cada idade modal, para cada ano de estudo, separado por sexo.

Ano	Unisex		Sexo Masculino		Sexo Feminino	
	Idade Modal	Qtde Expostos	Idade Modal	Qtde Expostos	Idade Modal	Qtde Expostos
2001	48	1662	48	1453	44	258
2002	49	1706	49	1490	45	265
2003	55	1723	55	1550	46	266
2004	55	1901	55	1718	47	275
2005	56	1955	56	1765	48	277
2006	55	2065	55	1812	49	283
2007	55	2145	55	1883	50	297
2008	55	2241	56	1941	51	308

Devido à disponibilidade dos dados da ELOS somente a partir de dezembro de 2003, restringimos os anos de estudo ao período de 2004 a 2008 para efeito de cálculo das taxas brutas de mortalidade e respectivas graduações, de forma a evitar resultados distorcidos, especificamente para o ano de 2003, no qual a quantidade de expostos aumentou excessivamente, porém não sendo possível a observação dos óbitos que ocorreram.

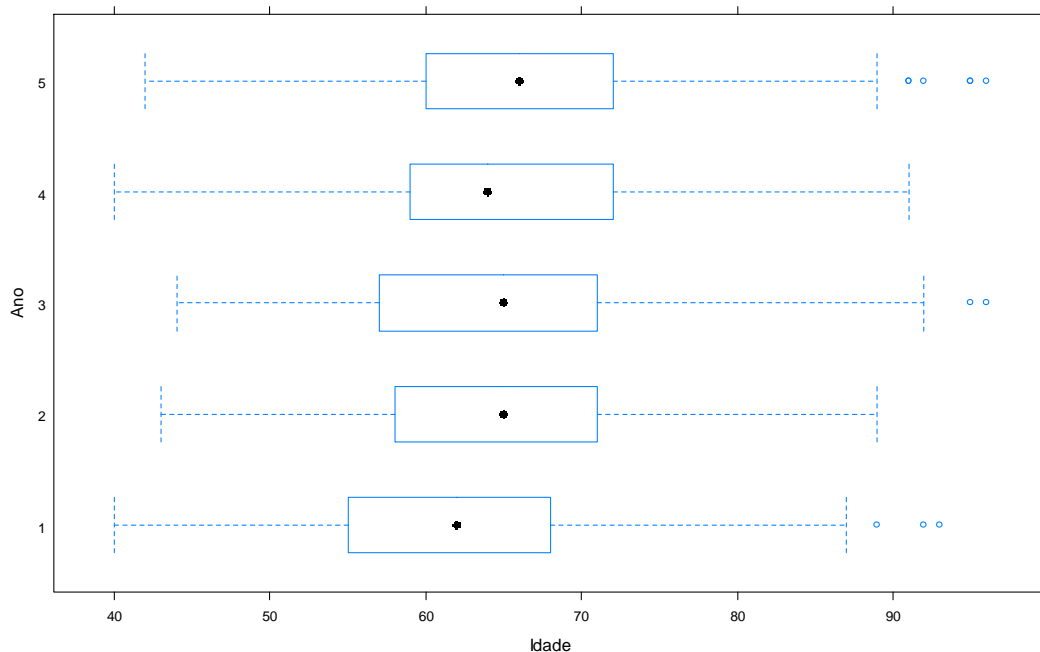


Figura 3.4 – Gráfico comparativo das idades de ocorrência de óbitos unisex quanto à dispersão, tendência central e medidas separatrizes, segundo o ano de estudo.

A Figura 3.4 resume as idades de ocorrência de óbitos, para os anos de 2004 a 2008, em que o box-plot da linha 1 corresponde ao ano de 2004, o da linha 2 corresponde ao ano de

2005 e assim sucessivamente. A Tabela 3.2 mostra que as idades de 55 e 56 anos são as que contêm maior número de expostos ao risco, tanto unissex como para o sexo masculino. Percebe-se da Figura 3.4 que entre as idades de 60 e 65 anos ocorre uma maior quantidade de mortes, sendo que a idade que apresenta maior proporção de óbitos tende a aumentar com o passar dos anos, conforme demonstra a Tabela 3.3 que aponta a evolução da idade média de ocorrência de óbitos, bem como outras medidas relacionadas a essas idades.

Tabela 3.3 - Resumo descritivo das idades de ocorrência de óbitos unissex, como medidas de tendência central e separatrizes, segundo o ano de estudo.

Medidas	Ano				
	2004	2005	2006	2007	2008
mínimo	40	44	40	42	53
1º quartil	55	58	57	58	61
moda	63	68	60	63	65
mediana	62	65	65	64	67
média	62,09	64,6	64,49	65,2	67,62
3º quartil	68	71	71	72	73
máximo	93	89	96	91	96

Extraí-se da Tabela 3.3 que no período compreendido entre 2004 e 2008, a faixa etária com maior exposição ao risco, é de aproximadamente 55 a 71 anos, correspondente à faixa etária compreendida ente o 1º e 3º quartis aproximadamente. Também, abaixo das idades correspondentes a 42 anos e acima de 90 anos, existem muito poucas mortes observadas. Desta forma, dada a baixa ocorrência de mortes nestas faixas etárias, optou-se por impor uma segunda restrição aos dados, graduando as taxas brutas de mortalidade apenas para as idades na faixa de 42 a 90 anos e, ainda assim, não se observou óbitos em algumas idades para os anos de estudo, conforme mostra a Tabela 3.4.

Tabela 3.4 – Idades nas quais não se observou óbitos em cada ano de estudo, considerando ambos os sexos.

	Anos de Estudo				
	2004	2005	2006	2007	2008
	51	48	48	48	48
Idades	76	84	54	49	51
sem	85	90	77	50	53
óbitos	86		88	70	87
	88		89	83	90
	90		90	90	

### 3.2 CÁLCULO DAS TAXAS BRUTAS DE MORTALIDADE

O cálculo das taxas brutas de mortalidade para cada ano de estudo e idade, é realizado de acordo com (2.8), dividindo o número de mortes observadas em cada idade  $x$  pela quantidade central de expostos ao risco. As taxas brutas de mortalidade unissex, considerando os anos de 2004 a 2008 e o intervalo de idade entre 42 e 90 anos, estão disponíveis no Anexo III. A Figura 3.5 apresenta o comportamento destas taxas para cada ano de estudo.

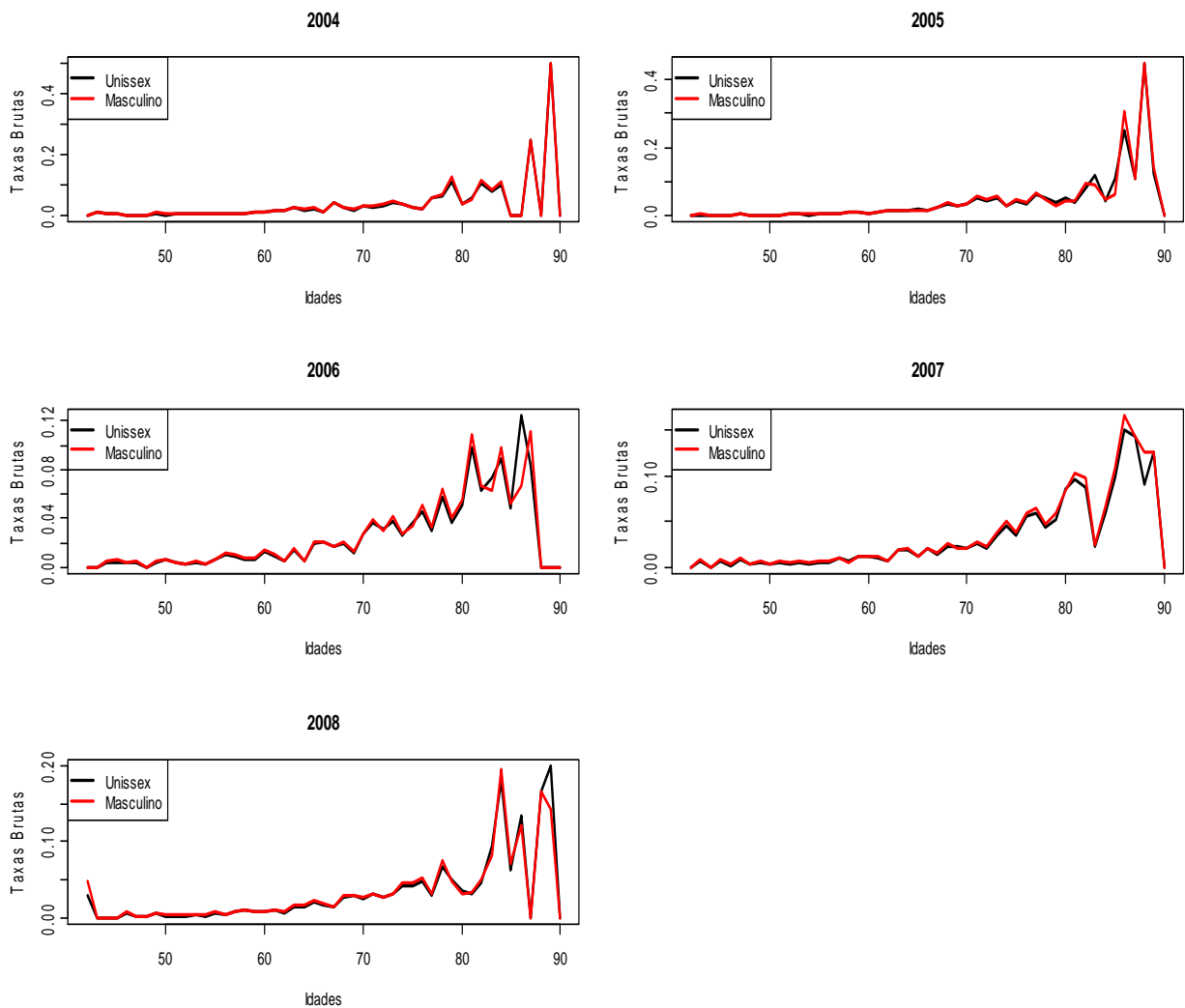


Figura 3.5 - Comportamento das taxas brutas de mortalidade por idade e sexo, para cada ano de estudo.

Como se pode observar, na Figura 3.5 os comportamentos das taxas brutas de mortalidade unissex e para o sexo masculino, são muito semelhantes, com exceção dos anos de 2006, 2007 e 2008 para as idades acima de 85 anos. Também, nota-se que em todos os anos de estudo, a faixa etária aproximada entre 42 e 75 anos, apresentam taxas brutas de

mortalidade com tendência crescente, exceto para o ano de 2008, no qual a idade de 42 anos apresenta taxa de mortalidade bem mais alta que as idades subseqüentes. Já as idades superiores á 75 anos – em todos os anos de estudo - apresentam comportamento bastante variável, devido à menor quantidade de expostos ao risco nesta faixa etária, de forma que os poucos óbitos observados provoca uma forte oscilação nas taxas brutas.

De forma geral, percebe-se que o comportamento das taxas brutas unissex é quase idêntico às taxas brutas para o sexo masculino. Assim, optou-se por graduar as taxas brutas considerando unissex.

### **3.3 TESTE DE ADERÊNCIA**

Conforme já mencionado na introdução deste estudo, a tábua de mortalidade é um instrumento fundamental no cálculo atuarial, uma vez que contém as probabilidades anuais de morte (e vida) teóricas de certa população, cujo objetivo é representar a real mortalidade de tal população, ou se aproximar bastante da mesma. Através da tábua de mortalidade, são dimensionadas rendas de aposentadorias e pensões, provisões matemáticas (recursos necessários ao pagamento de tais rendas ou indenização em seguros de vida), prêmios de seguros, contribuições e etc. Daí a importância da adequabilidade (aderência) da tábua à população de expostos ao risco, visando preservar o equilíbrio e solvência dos sistemas de previdência e seguridade.

A preocupação com a aderência é devida não só à constante evolução dos níveis de mortalidade, mas também com o fato de serem adotadas tábuas de mortalidade estrangeiras nos sistemas de previdência e seguridade, que é caso das tábuas AT's, baseadas na experiência americana. Logo, anteriormente à construção da tábua de mortalidade bayesiana, avaliaremos a aderência das tábuas AT-1949, AT-1983 e AT-2000 à população de expostos ao risco neste estudo, tendo em vista que as duas últimas tábuas são estabelecidas como parâmetros de mortalidade através das Resoluções CGPC N° 18 e CGPC N° 26.

Para tal avaliação, será empregado o teste das Reservas (Provisões) Matemáticas Efetivas versus Reservas (Provisões) Matemáticas Esperadas, proposto por Montello (2009). Este procedimento consiste em comparar, ao longo de  $n$  anos, a evolução da Reserva Matemática Efetiva com a evolução da Reserva Matemática Esperada de uma determinada população de aposentados, não decorrentes de invalidez, existente no final de um determinado ano  $t$ . Tanto a Reserva Matemática Efetiva como a Esperada, são calculadas com base numa determinada tábua de mortalidade, objeto do teste de aderência. Segundo este procedimento, a

tábua é considerada aderente, se ao final dos  $n$  anos de avaliação, a Reserva Matemática Esperada for igual ou superior a Reserva Matemática Efetiva.

Na população em estudo, cada exposto ao risco é um participante aposentado de alguma das EFPC's. Logo, seja  $x$  a idade de cada participante no momento inicial  $t=0$  (ano  $t$ ),  $B_x$  o valor do benefício de cada participante no momento  $t=0$  e  $n$  o período decorrido desde o momento inicial  $t=0$ . Assim, tem-se:

$$R_t = (13 \times B_x) \times a_{x+t}^{(12)}, \text{ com } t = 0, 1, 2, \dots, n$$

e

$$\hat{R}_t = (13 \times B_x) \times \frac{l_{x+t}}{l_x} \times a_{x+t}^{(12)}, \text{ com } t = 1, 2, \dots, n$$

em que  $R_t$  e  $\hat{R}_t$  são respectivamente as Reservas Matemáticas Efetivas e Esperadas calculadas no momento  $t$ ,  $(13 \times B_x)$  é o valor anualizado do benefício de cada participante já considerando o pagamento do 13º salário, a anuidade  $a_x^{(12)}$  e  $l_x$  já foram definidos anteriormente. Observe que  $l_{x+t}/l_x = {}_t p_x$ , isto é, a probabilidade de um participante com idade  $x$  no momento inicial  $t=0$ , atingir a idade  $x+t$ , com  $t > 0$ , a qual é utilizada no cálculo da Reserva Matemática Esperada. Portanto, o índice de aderência da tábua de mortalidade no momento  $t$ , é dado pela razão entre os totais das Reservas Matemáticas Esperadas e Efetivas, através da seguinte expressão:

$$A_d^t = \frac{\sum \hat{R}_t}{\sum R_t}$$

Logo, se  $A_d^t \geq 1$ , a tábua de mortalidade dá cobertura à sobrevivência ocorrida entre o momento inicial  $t=0$  e o momento  $t > 0$ , ao passo que, se  $A_d^t < 1$ , a tábua de mortalidade não dá cobertura à sobrevivência ocorrida no mesmo período.

Esta metodologia foi aplicada à população em estudo para se verificar a aderência das tábuas de mortalidade AT-1949, AT-1983 e AT-2000. Todavia, devido ao caráter confidencial dos valores dos benefícios dos participantes, os mesmos não foram disponibilizados, de forma que o valor de cada benefício foi considerado como sendo de R\$



1,00. Para o cálculo da anuidade  $a_x^{(12)}$  utilizada nos testes, adotou-se a taxa de desconto atuarial  $j=6\%$  ao ano. Quanto ao período observado para a realização do teste, tomou-se o momento inicial  $t=0$  como sendo 31/12/2003, com  $n=5$  anos, de forma que o momento final  $n+t$  é 31/12/2008.

As tabelas 3.5, 3.6 e 3.7 apresentam os resultados.

Tabela 3.5 – Teste de aderência da tábua de mortalidade AT-1949 à população exposta ao risco.

$t$	Reservas Efetivas	Reservas Esperadas	$A'_t$
2003	5.432.607,48		
2004	5.270.549,18	5.263.050,77	0,9986
2005	5.116.370,18	5.090.565,73	0,9950
2006	4.967.486,00	4.915.571,05	0,9895
2007	4.810.742,14	4.738.216,82	0,9849
2008	4.654.294,05	4.558.908,96	0,9795

Nota-se da Tabela 3.5 que os índices de aderência foram todos inferiores a 1. Logo, a tábua de mortalidade AT-1949 não dá cobertura à sobrevivência ocorrida no período em questão, se mostrando não aderente à mortalidade da população exposta ao risco e, portanto, não apropriada.

Tabela 3.6 – Teste de aderência da tábua de mortalidade AT-1983 à população exposta ao risco.

$t$	Reservas Efetivas	Reservas Esperadas	$A'_t$
2003	5.992.319,89		
2004	5.832.988,90	5.854.778,32	1,0037
2005	5.681.229,04	5.713.306,02	1,0056
2006	5.534.734,62	5.567.912,72	1,0060
2007	5.378.347,46	5.418.881,77	1,0075
2008	5.221.018,53	5.266.107,50	1,0086

Os índices de aderência contidos na Tabela 3.6 foram todos superiores a 1. Assim, a tábua de mortalidade AT-1983 dá cobertura à sobrevivência ocorrida no período em questão, se mostrando aderente à mortalidade da população exposta ao risco, sendo considerada apropriada. Logo, considerando a população em estudo, o resultado do teste confirma o

disposto na Resolução CGPC N° 18, a qual determina a adoção da tábua AT-1983 como parâmetro mínimo de mortalidade geral nas EFPC's.

Tabela 3.7 – Teste de aderência da tábua de mortalidade AT-2000 à população exposta ao risco.

$t$	Reservas Efetivas	Reservas Esperadas	$A_d^t$
2003	6.221.689,81		
2004	6.064.254,08	6.097.362,54	1,0055
2005	5.914.420,78	5.968.952,72	1,0092
2006	5.770.073,96	5.836.523,64	1,0115
2007	5.615.149,94	5.699.992,42	1,0151
2008	5.458.922,22	5.559.500,05	1,0184

Uma vez mostrada a aderência da tábua AT-1983, conforme a Tabela 3.6, não haveria necessidade de realização do teste de aderência da tábua AT-2000, uma vez que esta é mais conservadora que a primeira, implicando também em sua aderência à população exposta ao risco. Todavia, ainda assim foi realizado o teste para esta última tábua, que, conforme a Tabela 3.7, mostra que os índices de aderência foram todos superiores a 1 e, como já esperado, também foram superiores àqueles índices contidos na Tabela 3.6, para cada ano  $t$  observado. Logo, a tábua de mortalidade AT-2000 dá cobertura maior ainda à sobrevivência ocorrida no período em questão, apesar de que a tábua AT-1983 (menos conservadora) já se mostrou adequada.

Com base nos resultados dos testes realizados, verificou-se que a tábua de mortalidade AT-1983 é apropriada para dar cobertura ao risco de sobrevivência dos expostos ao risco. Todavia, nas próximas seções, serão abordados os modelos de graduação bayesiana de taxas de mortalidade, escolhendo-se aquele que melhor se adequar aos dados para a construção da tábua de mortalidade, cujas probabilidades de morte serão comparadas às probabilidades provenientes das três tábuas objeto dos testes de aderência.

### 3.4 GRADUAÇÃO BAYESIANA DAS TAXAS BRUTAS DE MORTALIDADE

Ainda que a faixa etária de 42 a 75 anos, em todos os anos de estudo, apresente uma tendência crescente das taxas brutas de mortalidade, conforme a mostra a Figura 3.5, este comportamento é bastante errático, não podendo ser considerado monotônico crescente em

relação à idade, como era de se esperar. Para a regularização deste comportamento não monotônico das taxas brutas em relação à idade, de forma a obter a tábua de mortalidade baseada na experiência da população exposta ao risco, é necessário um processo de suavização das mesmas, que nada mais é do que a graduação de tais taxas. Este processo, que visa um comportamento monotônico crescente, também propicia a graduação de taxas de mortalidade para idades cuja informação é insuficiente, porém isto dependerá do modelo utilizado na graduação, que neste trabalho, se dá em função da força de mortalidade  $\mu_x$  e utiliza a quantidade central de expostos ao risco  $e_x$ .

As taxas brutas graduadas são as probabilidades anuais de morte  $q_x$ , as quais compõem a tábua de mortalidade. A partir de certa idade, o comportamento dessas probabilidades será monotonicamente crescente, dado que a mortalidade humana se comporta assim a partir da mesma.

#### 3.4.1 MODELOS DE GRADUAÇÃO BAYESIANA

Os modelos bayesianos de graduação utilizados neste trabalho são aqueles propostos em Neves (2005), sendo dois locais e dois globais. O objetivo aqui é graduar as taxas brutas de mortalidade referentes à cobertura de sobrevivência nos planos de previdência complementar fechados, isto é, das EFPC's, baseando-se na experiência daquelas que disponibilizaram seus dados. Os modelos são implementados por algoritmos computacionalmente intensivos através de simulação estocástica de Monte Carlo via Cadeias de Markov (MCMC).

Como a graduação se dará em função da força de mortalidade, as mesmas podem ser obtidas a partir das tábuas biométricas usualmente adotadas nas EFPC's, aplicando-se a relação (2.7). Também, para tais tábuas, as probabilidades de morte apresentam o comportamento monotônico crescente a partir de aproximadamente 8 anos de idade. Logo, as forças de mortalidade também apresentam tal comportamento a partir de tal idade.

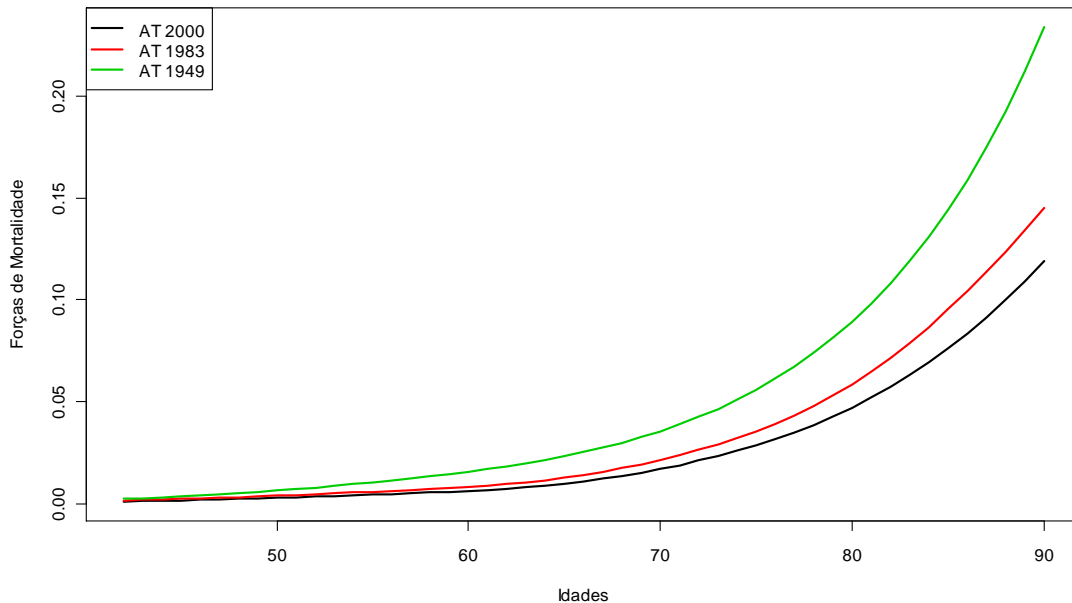


Figura 3.6: Comportamento das forças de mortalidade para as idades de 42 a 90 anos, segundo as tábuas biométricas AT-2000 básica, AT-1983 e AT-1949, todas para o sexo masculino.

A Figura 3.6 mostra as forças de mortalidade preconizadas por três tábuas biométricas. Como era de se esperar em todos os casos, o comportamento das forças de mortalidade, considerando a faixa etária especificada, é monotônico crescente. Nota-se que a tábua AT-1949 é a que apresenta maiores forças de mortalidade, conseqüentemente maiores probabilidades de morte. Tal tábua já não pode mais ser adotada como hipótese de mortalidade geral nas EFPC's, por força da Resolução CGPC N° 18 de 28/03/06, a qual estabelece como parâmetro mínimo de mortalidade geral, a adoção da tábua AT-1983. Também, o teste de aderência realizado na seção 3.3 mostrou que a tábua AT-1949 não é aderente à população em estudo. Todavia, a mesma tem sido adotada como parâmetro de mortalidade de inválidos em algumas EFPC's.

Nos modelos propostos para o processo de graduação, as forças de mortalidade futuras para um próximo período de tempo,  $\theta_i^r$ , para as classes de idade  $i=1, \dots, 49$  com  $r$  sendo a  $r$ -ésima iteração por MCMC; serão estimadas em função da distribuição preditiva do número de mortes  $d_i^r | \theta_i$ , onde é previsto o comportamento futuro desta quantidade, incorporando incertezas associadas aos parâmetros não observáveis  $\theta_i$ . O mesmo ocorre com as probabilidades de morte futuras  $q_{x+i-1}^r$  (NEVES, 2005).

Uma vez obtidas as estimativas das forças de mortalidade futuras  $\theta_i^r$ , aplica-se a simplificação (2.7) na obtenção das distribuições condicionais completas das probabilidades de morte futuras em cada idade, isto é,  $q_x, q_{x+1}, \dots, q_{x+k-1}$ , para as idades  $x, x+1, \dots, x+k-1$  onde  $x=42, k=49$ . Estas probabilidades são o objetivo final do processo de graduação. Assim, assumindo a quantidade central de expostos ao risco observada na idade  $x+i-1$  ( $e_i$ ) como constantes conhecidas e as iterações por MCMC, pode-se obter as estimativas das probabilidades de morte nas respectivas idades, da seguinte forma:

$$d_i^r | \theta_i \sim \text{Poisson}(e_i \cdot \theta_i),$$

em que  $\theta_i^r = \frac{d_i^r}{e_i}$  e  $q_{x+i-1}^r = 1 - \exp(-\theta_i^r)$ , para  $i = 1, \dots, 49$  e  $x = 42$  (3.1)

### 3.4.1.1 MODELO DE CARLIN MODIFICADO COM $\beta$ FIXO

O modelo de Carlin (1992), o qual chamaremos modelo 1, é uma adaptação do modelo de Boffit (1988) proposto em Neves (2005), com a restrição  $\theta \in R_I = \{\theta : 0 < \theta_1 < \theta_2 < \dots < \theta_{49} < \beta\}$  imposta ao vetor de parâmetros pertencente ao subconjunto  $\Omega = \{\theta; 0 < \theta_i < \infty, i = 1, \dots, 49\}$  de espaço euclidiano  $k$ -dimensional. Aqui,  $\beta$  é o limite máximo para a força de mortalidade no intervalo de tempo correspondente à idade de 90 anos, para a qual  $\beta=1$ .

O modelo 1 é não paramétrico, ou seja, não há o ajuste das taxas brutas de mortalidade a uma função de sobrevivência baseada em alguma lei de mortalidade. Desta forma, assumindo que  $d_i | \theta_i$  tem distribuição Poisson ( $e_i \cdot \theta_i$ ), considerando  $e_i$  e  $\theta_i$  constantes conhecidas, tem-se a seguinte função de verossimilhança:

$$l(\theta | d_i) \propto \prod_{i=1}^{49} \theta_i^{d_i} \exp(-e_i \cdot \theta_i) \quad (3.2)$$

em que  $d_i$  já foi definidos anteriormente e  $i=1, \dots, 49$  representa as classes de idade, com cada idade dada por  $x+i-1$ , iniciando em  $x=42$ . Também, é assumido que os  $\theta_i$ 's são independentes e identicamente distribuídos segundo uma Gama( $a, b$ ) restrita ao subconjunto  $R_I$ , com  $a=0,001$  e  $b=0,001$ . Assim, a distribuição a priori é dada por:

$$f(\theta | a, b) \propto \prod_{i=1}^{49} \theta_i^{a-1} \exp(-b \cdot \theta_i) I_{R_I}(\theta)$$

em que  $R_I = \{\theta : 0 < \theta_1 < \theta_2 < \dots < \theta_{49} < 1\}$ ,  $I_{R_I}(\theta) = 1$ , se  $\theta \in R_I$  e  $I_{R_I}(\theta) = 0$  caso contrário, com  $i=1, \dots, 49$ .

A função indicadora faz com que seja atendida a restrição imposta, conseqüentemente o sistema considera apenas os valores amostrados que pertençam aos intervalos especificados. Note que para este modelo não se utilizou hiperparâmetros desconhecidos, isto se deve às limitações tanto no pacote JAGS como no WinBUGS, uma vez que a utilização da função indicadora determinando limites quando os hiperparâmetros são desconhecidos, implica em construções inapropriadas nestes softwares.

Uma vez conhecida a distribuição a *priori* e a verossimilhança, obtém-se a distribuição conjunta dos parâmetros e dos dados através de:

$$p(d_i, \theta | a, b) = p(d_i | \theta) f(\theta | a, b) I_{R_I}(\theta)$$

em que  $I_{R_I}(\theta) = 1$ , se  $\theta \in R_I$  e  $I_{R_I}(\theta) = 0$  caso contrário. Simplificando a expressão, tem-se:

$$p(d_i, \theta | a, b) \propto \prod_{i=1}^{49} \theta_i^{d_i} \exp(-e_i \cdot \theta_i) \prod_{i=1}^{49} \theta_i^{a-1} \exp(-b \cdot \theta_i) I_{(\theta-i, \theta+i)}(\theta_i)$$

A partir da distribuição conjunta e aplicando-se o Teorema de Bayes, tem-se a distribuição a *posteriori* de  $\theta$ , dada por:

$$p(\theta | d_i, a, b) = \frac{p(d_i, \theta | a, b)}{p(d_i | a, b)} = \frac{p(d_i, \theta) f(\theta | a, b) I_{R_I}(\theta)}{\int p(d_i, \theta) f(\theta | a, b) I_{R_I}(\theta) d\theta}$$

Assim, a distribuição a *posteriori* condicional do parâmetro de interesse  $\theta$ , é obtida por:

$$p(\theta_i | d_i, a, b) = \frac{p(d_i, \theta) f(\theta | a, b) I_{(\theta_{i-1}, \theta_{i+1})}(\theta_i)}{\int p(d_i, \theta) f(\theta | a, b) I_{(\theta_{i-1}, \theta_{i+1})}(\theta_i) d\theta_i}$$

Por fim, dada a conjugação da distribuição a *priori* Gama do parâmetro  $\theta_i$  com a verossimilhança Poisson, temos a distribuição a *posteriori* condicional completa de  $\theta_i$ , considerando todos os demais parâmetros, dada por:

$$p(\theta_i | d_i, a, b) \propto \text{Gama}(\theta_i | a_i^*, b_i^*) I_{(\theta_{i+1}, \theta_{i-1})}(\theta_i)$$

em que  $a_i^* = a + c_i$ ,  $b_i^* = b + d_i$  e  $i=1, \dots, 49$ . Logo, a estrutura hierárquica bayesiana do modelo 1, é dada por:

$$\begin{aligned} d_i | \theta_i &\sim \text{Poisson}(e_i \cdot \theta_i), \text{ com } e_i \text{ conhecidos e } i = 1, \dots, 49 \\ \theta_i | a, b &\sim \text{Gama}(a, b) I_{(\theta_{i+1}, \theta_{i-1})}(\theta_i), \text{ com } \theta_0 = 0; a = b = 0,001 \end{aligned}$$

com predição realizada através de (3.1).

Para a implementação deste modelo por MCMC através do Amostrador de Gibbs - tanto no JAGS como no WinBUGS - foram simuladas três cadeias em paralelo com valores iniciais dos  $\theta_i$ 's distintos, visando evitar que as amostras geradas se concentrem em regiões em torno de uma moda local, em caso de multimodalidade da *posteriori* (NEVES, 2005). Nas simulações, os valores iniciais dos  $\theta_i$ 's em cada cadeia, são as forças de mortalidade extraídas das tábuas de mortalidade AT-2000, AT-1983 e AT-1949, aplicando-se a relação (2.6). Também, adotou-se o tamanho de amostras geradas para cada parâmetro desconhecido, como sendo de 120.000, através das três cadeias em paralelo, aproveitando uma a cada seis simulações (*thinning*) e considerando as primeiras 60.000 como *burn-in*. Assim, o número de iterações considerado de fato, é de 20.000, das quais as 10.000 primeiras são desprezadas (*burn-in*), sendo as inferências a *posteriori* feitas sobre as últimas 10.000. O computador utilizado na implementação dos modelos, foi um PC com processador Pentium 4 da Intel com 2.8GHz e 448MB de memória RAM, o qual levou aproximadamente 40 segundos para a implementação do modelo 1 no WinBUGS e em torno de 56 segundos para a implementação do mesmo modelo no JAGS.

Como diagnóstico da convergência dos parâmetros de interesse, são utilizadas três técnicas informais de análise gráfica: densidade, valores da estatística  $\hat{R}$  (R-hat - disponível

no WinBUGS) e os traços das cadeias. As Figuras 3.7 a 3.9 apresentam estes dispositivos, referentes à implementação do modelo 1 no WinBUGS e no JAGS, considerando o ano de estudo de 2008.

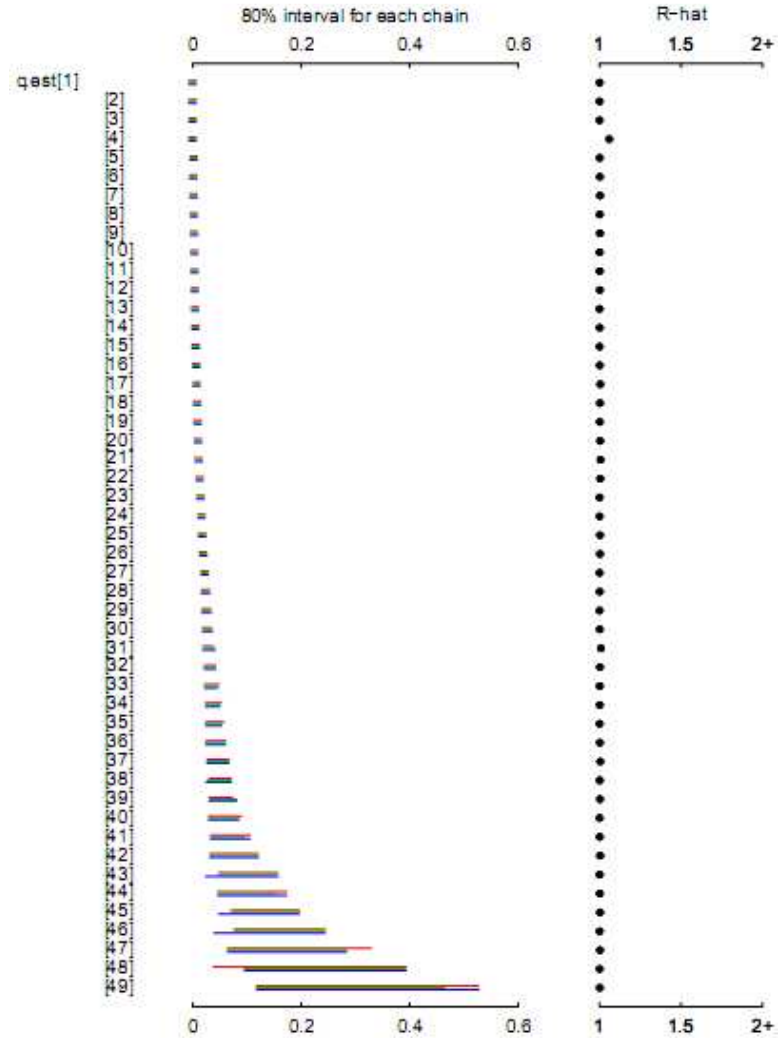


Figura 3.7 – Intervalos de credibilidade empíricos de 80% para as cadeias simuladas paralelamente e valores da estatística  $\hat{R}$  para análise da convergência das taxas graduadas pelo modelo 1 no WinBUGS, para o ano de 2008.

A Figura 3.7, mostra os comportamentos dos intervalos de credibilidade empíricos de 80% e da estatística  $\hat{R}$ , em que a variável “q.est” é o parâmetro de interesse sendo monitorado, isto é, as taxas de mortalidade graduadas (probabilidades de morte nas idades  $x - q_x$ ). Aqui, q.est[i] é a probabilidade de morte estimada  $\hat{q}_x$  para a idade  $x=42+i-1$ , com  $i=1,\dots,49$  sendo as classes de idade. Os intervalos empíricos de credibilidade de 80% para cada parâmetro  $q_x$ , são bem estreitos nas idades mais jovens, se concentrando em torno de valores bem baixos de  $q_x$ , se tornando mais amplos a partir de 84 anos (q.est[43]), se



concentrando em torno de valores mais altos de  $q_x$ . Este comportamento é devido ao aumento do erro padrão da estimativa à medida que se aumenta a idade, conseqüência de não haver quantidade suficiente de expostos ao risco e mortes observadas nas mesmas. Ainda, nos intervalos empíricos a cor azul se refere aos parâmetros de interesse  $q_x$ , considerando as três cadeias em conjunto ( $R_p$ ), ao passo que a cor verde se refere à média nas três cadeias destes intervalos ( $R_w$ ) para estes parâmetros. A Figura 3.8 permite uma melhor visualização do comportamento das taxas de mortalidade graduadas, onde os pontos azuis e verdes são correspondentes a  $R_p$  e  $R_w$  já descritos e o ponto vermelho é a estimativa pontual da probabilidade de morte na respectiva idade.

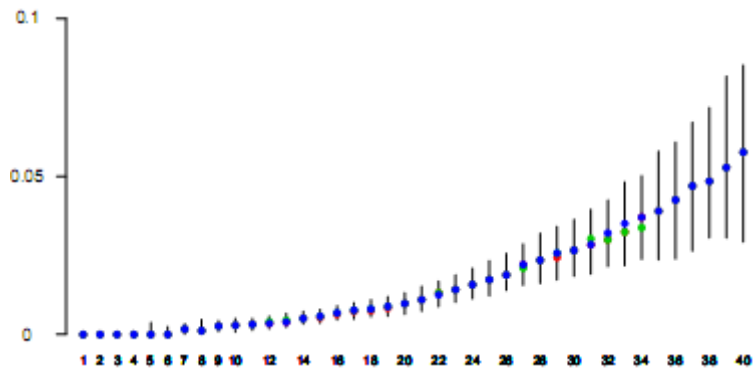


Figura 3.8 - Intervalos de credibilidade empíricos de 80% para as taxas de mortalidade graduadas através do modelo 1 para o ano de 2008, considerando as três cadeias simuladas em conjunto, no pacote WinBUGS.

A estatística  $\hat{R} = R_p / R_w$ , está associada às inferências *a posteriori* realizadas sobre a distribuição dos parâmetros de interesse. Quanto mais próximo de 1, melhor as cadeias se misturam, indicando convergência das cadeias para os parâmetros de interesse. Assim, com base nesta estatística e de acordo com a Figura 3.7, há indicativo de convergência dos parâmetros de interesse em todas as idades para as quais foram graduadas as taxas de mortalidade, considerando os dados referentes ao ano de 2008.

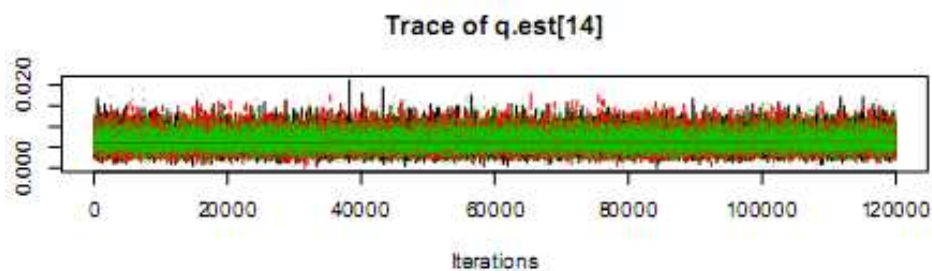


Figura 3.9 – Comportamento das cadeias (traços) simuladas para  $\hat{q}_{55}$ , considerando as últimas 10.000 iterações, para o ano de 2008.

Na Figura 3.9, inclusa a título de exemplificação, observa-se que as cadeias paralelas se misturam bem, oscilando em torno de um valor médio, convergindo para a estimativa  $\hat{q}_{55} = 0,00667$ . Também a título de exemplificação, a Figura 3.10 evidencia que  $\hat{q}_{55}$  é uma boa estimativa para  $q_{55}$ , devido à suavidade da densidade *a posteriori* obtida por estimação de densidade a partir das amostras, a qual não apresenta multimodalidade. A estimativa pontual  $\hat{q}_{55}$  é obtida através da média das observações simuladas da *posteriori*.

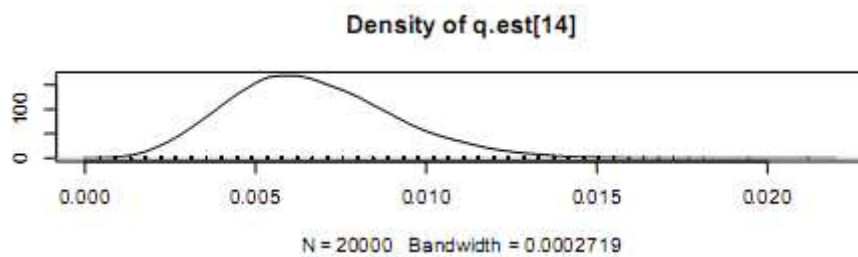


Figura 3.10 – Densidade *a posteriori* de  $\hat{q}_{55}$ , correspondente a suavização do histograma através do método *kernel*, para o ano de 2008.

A densidade acima foi extraída do pacote JAGS, devido ao fato do WinBUGS disponibilizar densidades somente para cada cadeia individualmente, enquanto o JAGS o faz considerando as cadeias em conjunto.

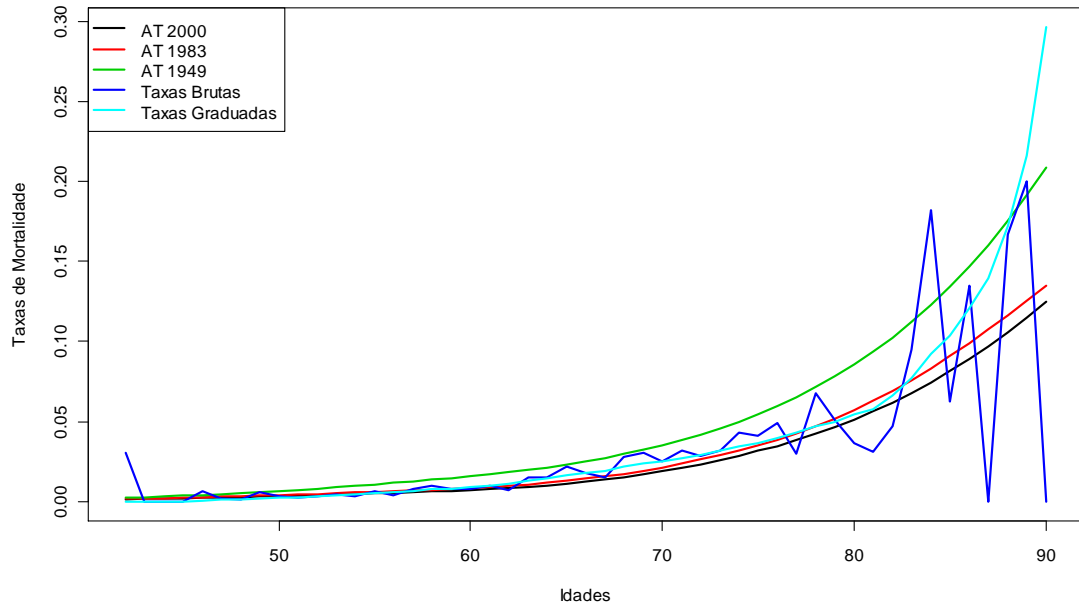


Figura 3.11 – Comparativo das taxas de mortalidades das tábuas AT-2000, AT-1983 e AT-1949 com as taxas brutas de mortalidade e taxas graduadas através do modelo 1, considerando o ano de 2008.

O comportamento das taxas graduadas na Figura 3.11, permanece bem próximo das taxas de mortalidade da tábua AT-1983 até aproximadamente 83 anos de idade. Todavia, a partir desta idade, apresenta um crescimento acentuado, porém se mantendo abaixo das taxas de mortalidade da tábua AT-1949, até aproximadamente 88 anos. Apesar deste crescimento acentuado nas idades mencionadas, as taxas de mortalidade graduadas apresentaram a suavidade e o comportamento monotônico crescente esperado, aparentando um bom ajuste aos dados, ao menos até 83 anos de idade. O comportamento apresentado acima dos 83 anos é devido à baixa exposição ao risco a partir desta idade, bem como a não observação de mortes para algumas idades superiores a esta, conforme demonstra a Tabela 3.4. Este fato também justifica a grande amplitude dos intervalos de credibilidade apresentados na Figura 3.7, a partir de aproximadamente 84 anos, refletindo o aumento dos erros padrões das estimativas das probabilidades de morte.

### 3.4.1.2 MODELO COM REPARAMETRIZAÇÃO TIPO BROFFIT

Este, que chamaremos modelo 2, é também não paramétrico, evitando o uso da função indicadora e implementando uma reparametrização no modelo 1, diretamente no algoritmo, mantendo a mesma restrição  $\theta \in R_I = \{\theta : 0 < \theta_1 < \theta_2 < \dots < \theta_{49} < 1\}$  imposta ao vetor de

parâmetros (NEVES, 2005). A reparametrização é dada por  $\theta_i = \theta_{i-1} + \phi_i$ , com  $i=1, \dots, 49$  sendo as classes de idade. Portanto, pode-se determinar os  $\phi_i$ 's em função dos  $\theta_i$ 's, da seguinte forma:

$$\phi_i = \theta_i - \theta_{i-1}, \text{ com } \theta_0 = 0 \text{ e } i = 1, \dots, 49 \quad (3.3)$$

Para garantir a restrição do crescimento dos  $\theta_i$ 's no modelo, considera-se os  $\phi_i$ 's independentes e distribuídos segundo uma Gama( $a, b$ ), fixando  $a=0,8$  e  $b=231$ , obtidos do ajuste deste modelo aos dados. Os  $\phi_i$ 's são incrementos positivos dos componentes  $\theta_i$ 's, onde cada parâmetro  $\theta_i$  é formado pela soma do parâmetro anterior  $\theta_{i-1}$  com o incremento  $\phi_i$  gerado pela distribuição Gama. Portanto, considerando a função de verossimilhança dada em (3.2) e a distribuição conjunta dos  $\phi_i$ 's, dada por:

$$p(d_i, \phi | a, b) \propto \prod_{i=1}^{49} \theta_i^{d_i} \exp(-e_i \cdot \theta_i) \prod_{i=1}^{49} \phi_i^{a-1} \exp(-b \phi_i) \propto \prod_{i=1}^{49} \theta_i^{d_i} \phi_i^{a-1} \exp(-e_i \theta_i - b \phi_i)$$

em que  $\phi_i > 0$ ,  $\theta_i = \phi_1 + \dots + \phi_i$ , e  $i = 1, \dots, 49$ .

Portanto, a estrutura hierárquica do modelo 2, é dada por:

$$\begin{aligned} d_i | \theta_i &\sim \text{Poisson}(e_i \cdot \theta_i), \text{ com } e_i \text{ conhecidos e } i = 1, \dots, 49; \\ \phi_i | a, b &\sim \text{Gama}(a, b), \text{ com } a = 0,8 \text{ e } b = 231; \\ \theta_i &= \theta_{i-1} + \phi_i, \text{ sendo } \theta_0 = 0 \end{aligned}$$

com predição realizada através de (3.1).

Nas simulações, os valores iniciais dos  $\phi_i$ 's em cada cadeia, são determinados através de (3.3), a partir das forças de mortalidade das tábuas biométricas AT-2000, AT-1983 e AT-1949, aplicando-se a relação (2.6). Para este modelo, adotou-se também os mesmos tamanhos de amostras geradas, *burn-in* e *thinning* utilizados na implementação do modelo 1, simulando também as três cadeias em paralelo. O tempo gasto na implementação do modelo 2 no WinBUGS é de aproximadamente 55 minutos, por ser um modelo mais computacionalmente intensivo.

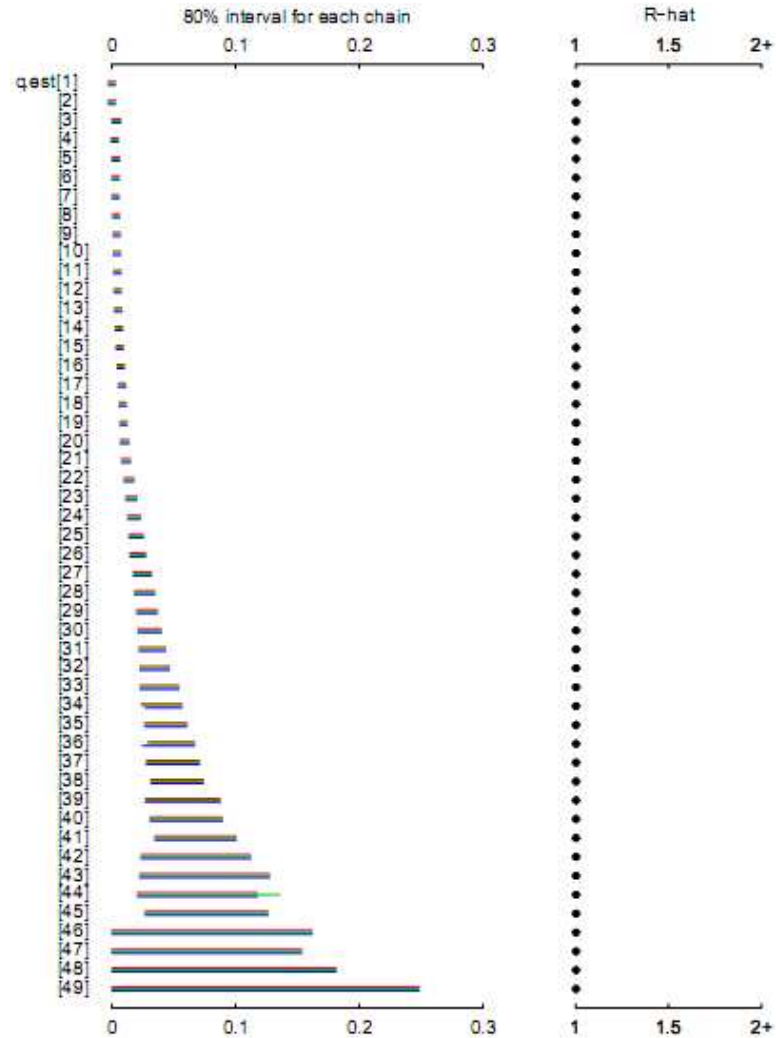


Figura 3.12 – Intervalos de credibilidade empíricos de 80% para as cadeias simuladas paralelamente e valores da estatística  $\hat{R}$  para análise da convergência das taxas graduadas pelo modelo 2 no WinBUGS, para o ano de 2008.

A Figura 3.12 apresenta os gráficos para análise de convergência dos parâmetros de interesse, referentes à implementação do modelo 2 no WinBUGS, considerando o ano de estudo de 2008. Os intervalos de credibilidade empíricos de 80% para cada  $q_x$  - considerando as cadeias simuladas paralelamente - são bem estreitos nas idades mais jovens, se concentrando em torno de valores bem baixos de  $q_x$ , se tornando mais amplos gradativamente conforme a idade avança. Assim como no modelo 1, este comportamento pode ser atribuído ao aumento do erro padrão da estimativa à medida que se aumenta a idade, pelos mesmos motivos já expostos.

Quanto aos valores da estatística  $\hat{R}$ , apresentam indicativo de que as cadeias estão se misturando bem, havendo a convergência dos parâmetros de interesse, dado que para todos os valores de  $\hat{q}_x$  ( $q.est[i]$ ) o valor de  $\hat{R}$  é igual a 1.

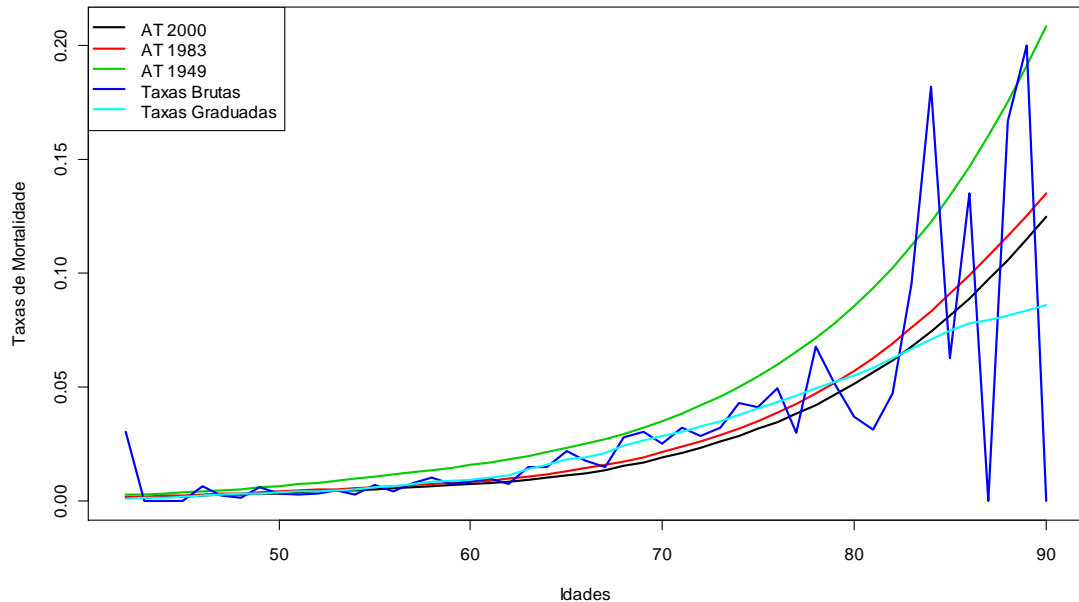


Figura 3.13 – Comparativo das taxas de mortalidades das tábuas AT-2000, AT-1983 e AT-1949 com as taxas brutas de mortalidade e taxas graduadas através do modelo 2, considerando o ano de 2008.

As estimativas deste modelo, de forma geral acompanham bem a tendência das taxas brutas de mortalidade, apresentando o comportamento monotônico crescente e a suavidade desejada para as taxas de mortalidade graduadas. Todavia, ao contrário do modelo 1, as taxas graduadas pelo modelo 2 se mostram mais conservadoras, de forma que a partir de aproximadamente 82 anos de idade, chegam a alcançar patamares inferiores às taxas de mortalidade da tábua AT-2000. Também, nota-se da Figura 3.13 que as taxas graduadas entre aproximadamente 62 e 79 anos de idade, não estão tão próximas das taxas de mortalidade da tábua AT-1983, sendo menos conservadoras. Portanto, exclusivamente com base na análise gráfica da Figura 3.13, o modelo 2 também parece proporcionar um bom ajuste aos dados, uma vez que acompanha a tendência das taxas brutas de mortalidade e induz um comportamento suavizado e monotônico, conforme esperado.

### 3.4.1.3 MODELO COM FORÇA DE MORTALIDADE TIPO GOMPERTZ

Este modelo, o qual chamaremos modelo 3, é paramétrico e pressupõe que a força de mortalidade obedece a uma função de sobrevivência de acordo com uma lei de mortalidade. Neste caso adotou-se a Lei de Gompertz, proposta em 1825 (NEVES, 2005). Esta lei prevê que a força de mortalidade cresce em progressão geométrica, conforme descrito em Sherris (1999), cuja expressão é dada por:

$$\mu_x = \beta c^x$$

em que  $\beta > 0$ ,  $c \geq 1$  e  $x \geq 0$ , sendo  $x$  a idade em anos. Como estamos considerando  $\mu_y = \theta_i$ ,  $x+i-1 \leq y \leq x+i$ ,  $i=1, \dots, 49$  e  $x=42$  anos, tem-se  $\theta_i = \beta c^{Z_i}$ , com  $Z_i = x+i-1$  que representa as idades em estudo e  $i$  representa as classes de idade.

Este modelo paramétrico assume a forma de um modelo linear generalizado (MLG) bayesiano, com base no modelo clássico não preditivo proposto por Haberman e Renshaw (1996), conforme descrito em Neves (2005). Como anteriormente, assume-se que o número de mortes observadas na idade  $x+i-1$  ( $d_i$ ) é distribuído segundo uma  $Poisson(e_i \theta_i)$ , onde  $e_i$  são constantes conhecidas e representam a quantidade central de expostos ao risco para a idade  $x+i-1$ , com  $i=1, \dots, 49$ . Logo,  $d_i | \theta_i$  são independentes e identicamente distribuídos, com  $E(d_i | \theta_i) = \lambda_i = e_i \theta_i$ .

Conforme Dobson (1994), a função de ligação canônica  $\eta_i$  da distribuição Poisson é dada por  $\log(\lambda_i)$ , seguindo a estrutura do MLG. Assim, a estrutura do modelo é dada por:

$$\eta_i = \log(\lambda_i) = \log(e_i) + \log(\beta) + Z_i \log(c)$$

em que  $i=1, \dots, 49$ . Portanto, o modelo possui dois parâmetros  $\beta$  e  $c$  a serem estimados e uma variável explicativa  $Z_i$ , sendo  $\log(e_i)$  o *offset* e  $\log(\beta)$  o intercepto.

No processo de inferência bayesiana, atribuem-se distribuições *a priori* normais aos parâmetros  $\beta$  e  $c$ , ressaltando que no processo de simulação no pacote WinBUGS, utiliza-se o parâmetro de precisão  $\tau = \sigma^{-2}$  ao invés da variância. Finalmente, a estrutura hierárquica bayesiana do modelo 3, é dada por:

$$\begin{aligned}
 d_i | \theta_i &\sim \text{Poisson}(\lambda_i), \lambda_i = e_i \cdot \theta_i, \text{ com } e_i \text{ conhecidos e } i = 1, \dots, 49; \\
 \log(\lambda_i) &= \log(e_i) + \log(\beta) + Z_i \log(c); \\
 \theta_i &= \beta c^{Z_i}; \\
 \beta &\sim \text{Normal}(0; 10000) I_{(0, \infty)}(\beta); \\
 c &\sim \text{Normal}(0; 10000) I_{(1, \infty)}(c)
 \end{aligned}$$

com predição realizada através de (3.1).

Diferente dos modelos anteriores, aqui os valores iniciais na implementação do modelo, são aqueles atribuídos aos parâmetros  $\beta$  e  $c$ , os quais são tratados como nós estocásticos no WinBUGS, sendo que as forças de mortalidades futuras  $\theta_i^t$  dadas em função destes, caracterizam nós lógicos. Também, aqui não se tem mais três cadeias sendo simuladas paralelamente, pois as forças de mortalidade são calculadas em função dos valores simulados dos parâmetros  $\beta$  e  $c$ , de forma que se tem uma só cadeia simulada. No demais, permanecem as mesmas condições consideradas nos modelos anteriores: mesmos tamanhos de amostras geradas, *burn-in* e *thinning* para a implementação.

Este modelo foi implementado apenas no WinBUGS, levando aproximadamente 45 segundos na graduação das taxas de mortalidade, no computador utilizado.

A Figura 3.14 apresenta o comportamento das taxas de mortalidade graduadas, juntamente com os respectivos intervalos de confiança empíricos de 80%, referentes à implementação do modelo 3, considerando o ano de estudo de 2008.

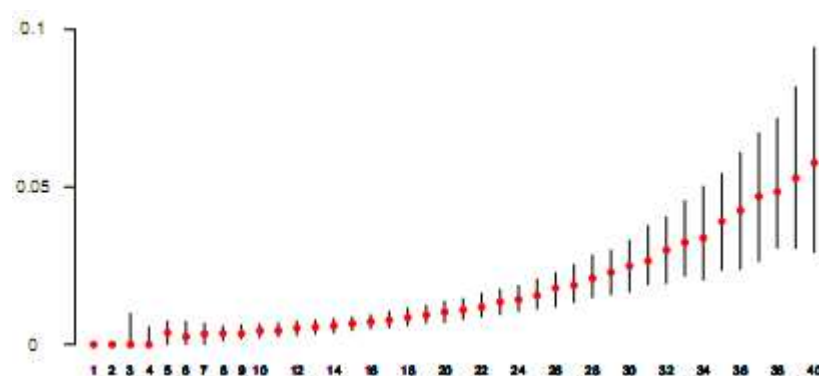


Figura 3.14 – Intervalos de credibilidade empíricos de 80% para as taxas de mortalidade graduadas, considerando uma cadeia simulada através do modelo 3 no WinBUGS.

Os intervalos de confiança empíricos de 80% para cada  $q_x$  - considerando apenas uma cadeia simulada - são bem estreitos nas idades mais jovens, se concentrando em torno de valores bem baixos de  $q_x$ , se tornando mais amplos gradativamente conforme a idade avança,



devido ao aumento dos erros padrões das estimativas  $\hat{q}_x$ . Apesar deste aumento dos erros à medida que a idade avança, no modelo 3 os intervalos de credibilidade não são tão amplos quanto nos modelos anteriores, aparentando uma menor incerteza em relação às probabilidades de morte estimadas.

Note que aqui não há valores da estatística  $\hat{R}$ , como para modelos anteriores por não haver mais de uma cadeia simulada. Neste caso, o diagnóstico de convergência é realizado através do traço da cadeia e pela densidade estimada, conforme Figuras 3.15 e 3.16 para  $q_{66}$  a título de exemplificação.

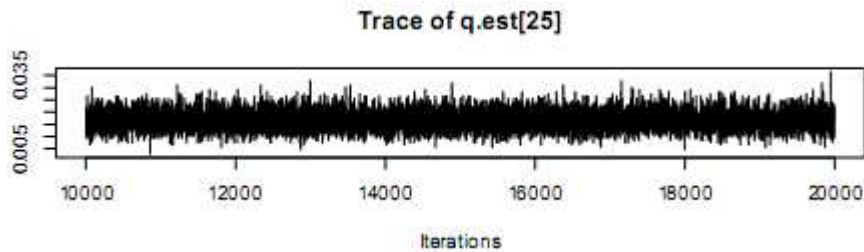


Figura 3.15 – Comportamento da cadeia (traço) simulada para  $\hat{q}_{66}$ , considerando as últimas 10.000 iterações, disponibilizado pelo pacote WinBUGS, para o ano de 2008.

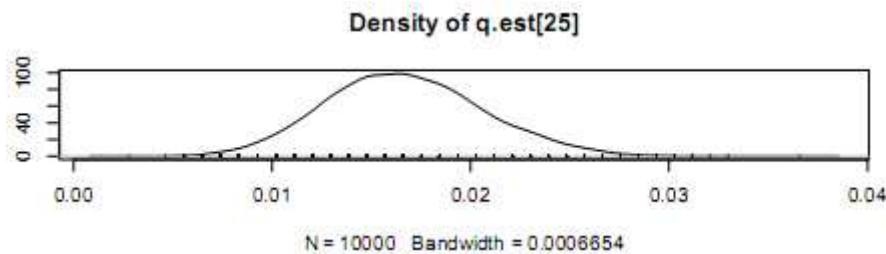


Figura 3.16 – Densidade *a posteriori* de  $\hat{q}_{66}$ , correspondente a suavização do histograma através do método *kernel*, disponibilizada pelo pacote WinBUGS, para o ano de 2008.

Como se pode observar, se analisado o *trace* da cadeia para a idade de 66 anos, a mesma parece estar oscilando em torno de um valor médio e, portanto, convergindo. Também, quando analisado o comportamento da densidade *a posteriori* de  $\hat{q}_{66}$ , nota-se que a mesma apresenta a suavidade desejada e ausência de multimodalidade, implicando que  $\hat{q}_{66}$  é um bom estimador para  $q_{66}$ .

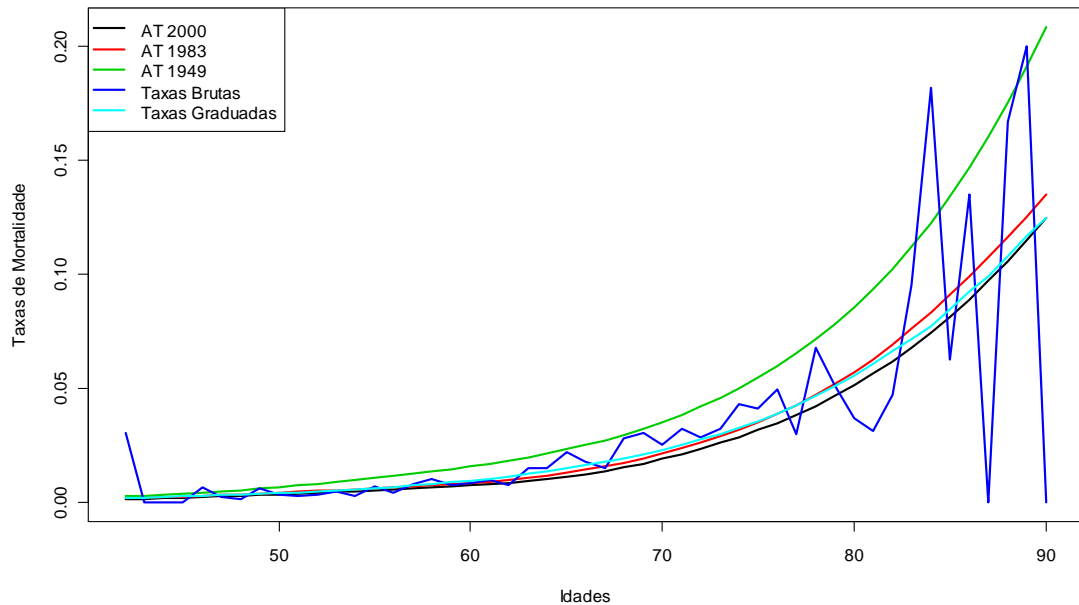


Figura 3.17 – Comparativo das taxas de mortalidades das tábuas AT-2000, AT-1983 e AT-1949 com as taxas brutas de mortalidade e taxas graduadas através do modelo 3, considerando o ano de 2008.

Da mesma forma que no modelo 2, as taxas graduadas através do modelo 3 apresentam comportamento monotônico crescente, porém as probabilidades de morte estimadas atingem patamares mais elevados à medida que a idade ultrapassa os 79 anos, sendo menos conservadoras que no primeiro. É notável a grande proximidade entre as taxas de mortalidade graduadas e aquelas provenientes da tábua AT-1983 para as idades entre 42 e aproximadamente 82 anos, se tornando bastante próximas da tábua AT-2000 a partir de 83 anos.

Portanto, exclusivamente com base na análise gráfica da Figura 3.17, este modelo paramétrico também parece apresentar um bom ajuste aos dados, uma vez que apresenta o comportamento monotônico crescente e parece acompanhar a tendência apresentada pelas taxas brutas.

#### 3.4.1.4 MODELO COM FORÇA DE MORTALIDADE TIPO MAKEHAM

Este modelo – chamado modelo 4 - também é paramétrico, diferindo do modelo 3 apenas por uma mudança na função de sobrevivência que se ajusta a força de mortalidade, isto é, outra lei de mortalidade, neste caso a Lei de Makeham. A Lei de Gompertz contemplou apenas as causas naturais de mortes, ao passo que em 1867, Makeham sugeriu a adição de

uma constante  $\alpha$  à expressão matemática da Lei de Gompertz, de forma a considerar também os óbitos ocorridos por causas acidentais (SHERRIS, 1999). A função desta constante é estimar aquela parcela da força de mortalidade que independe da idade (NEVES, 2005). Tal lei é representada pela seguinte expressão matemática:

$$\mu_x = \alpha + \beta c^x$$

em que  $\alpha \geq 0, \beta > 0, c \geq 1$  e  $x \geq 0$ , sendo  $x$  a idade em anos. Desta forma, temos que  $\theta_i = \alpha + \beta c^{Z_i}$ , com  $Z_i = x + i - 1, x=42$  e  $i=1, \dots, 49$  representando as classes de idade em estudo. Assim, a estrutura do modelo passa a ser:

$$\eta_i = \log(\lambda_i) = \log(e_i) + \log(\alpha + \beta c^{Z_i})$$

em que  $i=1, \dots, 49$ . Portanto, o modelo passa a ter três parâmetros a serem estimados:  $\alpha, \beta$  e  $c$ , mantendo a variável explicativa  $Z_i$ , sendo  $\log(e_i)$  o *offset*.

Para possibilitar o processo de inferência bayesiana, atribui-se distribuições *a priori* Normais não informativas aos parâmetros  $\alpha, \beta$  e  $c$ , ressaltando que no processo de simulação no pacote WinBUGS, utiliza-se o parâmetro de precisão  $\tau = \sigma^{-2}$  ao invés da variância. Assim, a estrutura hierárquica bayesiana do modelo 4, é dada por:

$$\begin{aligned} d_i | \theta_i &\sim \text{Poisson}(\lambda_i), \lambda_i = e_i \cdot \theta_i, \text{ com } e_i \text{ conhecidos e } i = 1, \dots, 49; \\ \log(\lambda_i) &= \log(e_i) + \log(\alpha + \beta c^{Z_i}); \\ \theta_i &= \alpha + \beta c^{Z_i} \\ \alpha &\sim \text{Normal}(0; 10000) I_{(0, \infty)}(\alpha); \\ \beta &\sim \text{Normal}(0; 10000) I_{(0, \infty)}(\beta); \\ c &\sim \text{Normal}(0; 10000) I_{(1, \infty)}(c) \end{aligned}$$

com predição realizada através de (3.1).

Em comparação com o modelo 3, há agora três valores iniciais na implementação, correspondentes aos parâmetros  $\alpha, \beta$  e  $c$ , tratados como nós estocásticos no WinBUGS, sendo as forças de mortalidades ( $\theta_i'$ ) classificadas como nós lógicos, dadas em função destes parâmetros. Da mesma forma que no modelo 3, é simulada apenas uma cadeia. Também,

permanecem as mesmas condições consideradas nos modelos anteriores: mesmos tamanhos de amostras geradas, *burn-in* e *thinning* para a implementação.

O tempo de implementação deste modelo no WinBUGS, foi de aproximadamente 55 segundos para a graduação das taxas de mortalidade, mantendo o mesmo computador.

A Figura 3.18 apresenta o comportamento das taxas de mortalidade graduadas, juntamente com os respectivos intervalos de confiança empíricos de 80%, referentes à implementação do modelo 4 no WinBUGS, considerando o ano de estudo de 2008.

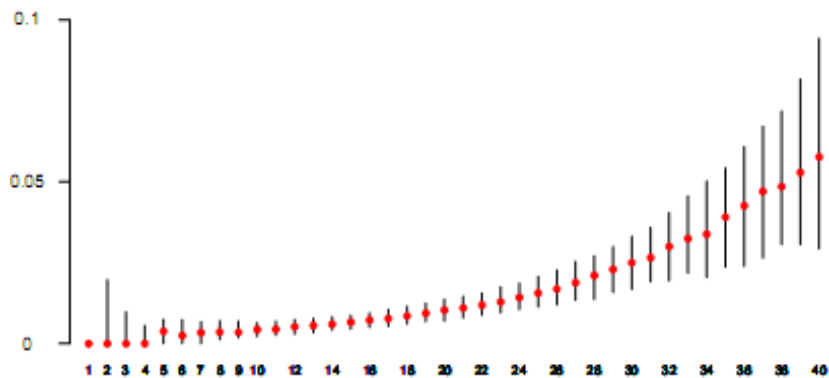


Figura 3.18 – Intervalos de credibilidade empíricos de 80% para as taxas de mortalidade graduadas, considerando uma cadeia simulada através do modelo 4 no WinBUGS.

Como se pode observar, o comportamento das taxas graduadas por este modelo é muito semelhante àsquelas do modelo 3. Logo, o comportamento dos intervalos de credibilidade empíricos de 80% para cada  $q_x$  também são muito semelhantes, possuindo a mesma interpretação. Apesar desta semelhança com o modelo 3, neste as taxas graduadas são um pouco mais altas, sendo um pouco menos conservadoras, como também mostra a Figura 3.19. Também, aqui não há valores da estatística  $\hat{R}$ , pelo mesmo motivo citado para o modelo 3, sendo o diagnóstico de convergência realizado através do *trace* da cadeia e pela densidade suavizada (*kernel*).

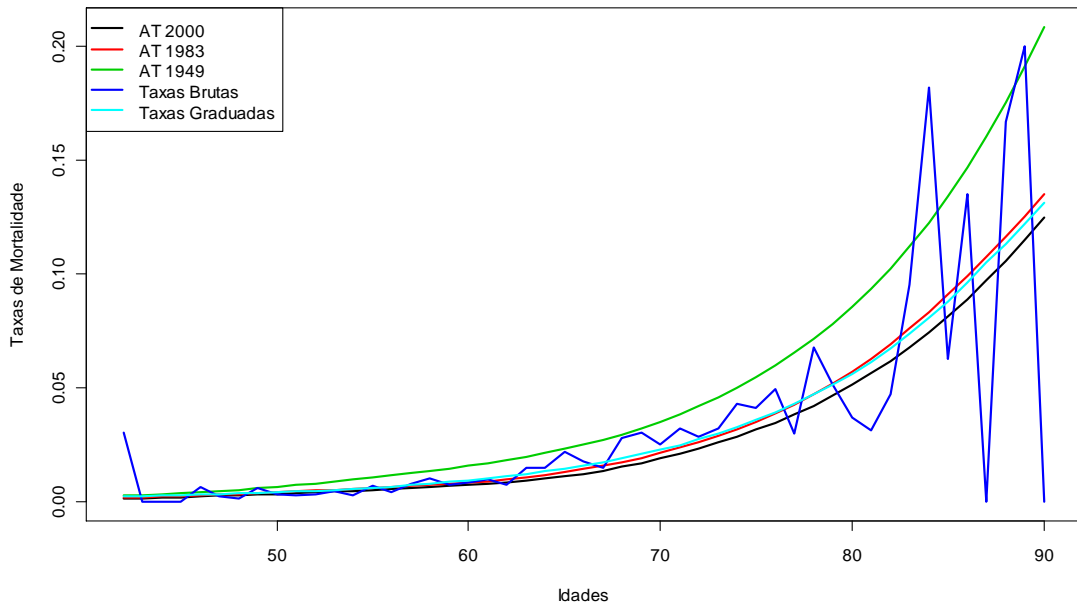


Figura 3.19 – Comparativo das taxas de mortalidades das tábuas AT-2000, AT-1983 e AT-1949 com as taxas brutas de mortalidade e taxas graduadas através do modelo 4, considerando o ano de 2008.

Da mesma forma que no modelo 3, neste as taxas graduadas também permanecem bastante próximas das taxas da tábua AT-1983, porém em todas as idades, sendo um pouco mais conservadoras a partir de aproximadamente 82 anos. Também, as taxas de mortalidade graduadas através do modelo 4 apresentam comportamento monotônico crescente, muito semelhante àquele apresentado pelas taxas graduadas através do modelo 3, porém aqui, as taxas são um pouco mais elevadas. Uma das causas disto é a inclusão do parâmetro  $\alpha$ , cuja finalidade é explicar as mortes ocorridas por causas acidentais, de acordo com a Lei de Makeham.

Logo, exclusivamente com base na análise gráfica da Figura 3.19, também se pode dizer que este modelo paramétrico apresenta um bom ajuste aos dados, pois além do comportamento monotônico crescente, parece acompanhar relativamente a tendência apresentada pelas taxas brutas.

### 3.4.2 COMPARAÇÃO ENTRE OS MODELOS DE GRADUAÇÃO PROPOSTOS

Para a escolha do modelo de graduação bayesiana mais adequado dentre os propostos, foi adotado o método DIC (*Deviance Information Criterion*), desenvolvido por Spiegelhalter (1997). Os valores das estatísticas que compõem o DIC são disponibilizados

pelo pacote WinBUGS, tratando-se de um critério de escolha do modelo. Este critério é uma generalização bayesiana do método AIC (*Akaike Information Criterion*), ou ainda, uma medida bayesiana de ajuste dos dados, penalizado por um termo adicional de complexidade, sendo utilizado para comparar modelos hierárquicos complexos, nos quais o número de parâmetros pode até mesmo ultrapassar o número de observações.

O método é composto de dois termos, onde o primeiro representa uma medida de ajustamento do modelo ( $\bar{D}$ ) e o segundo mede o número efetivo de parâmetros ( $pD$ ), assim:

$$DIC = \bar{D} + pD$$

sendo utilizada a média *a posteriori* da função desvio como medida bayesiana de ajustamento ou adequação de modelos, dada por:

$$\bar{D} = E_{\theta|y_0} [D(\theta)] = E_{\theta|y_0} [-2 \ln p(y_0 | \theta)]$$

A idéia aqui é de quanto maior for a verossimilhança, melhor o ajuste aos dados. Portanto, quanto menor o valor da estatística  $\bar{D}$ , melhor o ajuste do modelo aos dados. Quanto à estatística  $pD$ , é obtida pela diferença entre a média *a posteriori* da função desvio e a função desvio da média *a posteriori* dos parâmetros de interesse. Este cálculo fornece uma eficiente e direta aproximação bayesiana para definição do número efetivo de parâmetros, obtendo uma medida de complexidade para os modelos propostos. Assim, tem-se que a estatística  $pD$  é dada por:

$$\begin{aligned} pD &= \bar{D} - D(\bar{\theta}) = \bar{D} - \hat{D} = E_{\theta|y_0} [D(\theta)] - D(E_{\theta|y_0} [\theta]) \\ &= E_{\theta|y_0} [-2 \ln p(y_0 | \theta)] + 2 \ln p(y_0 | \bar{\theta}) \end{aligned}$$

Conforme descrito em Neves (2005),  $pD$  é sempre positivo para uma função de verossimilhança log-côncava, sendo negativo em duas situações:

- quando a função de verossimilhança não é log-côncava;
- quando a distribuição *a posteriori* para um parâmetro é bimodal simétrica, caso em que a média *a posteriori* é um estimador pobre.

Em suma, o DIC pode ser entendido como uma medida clássica de ajuste e adequação combinada com uma medida de complexidade do modelo. Assim, para análise dos modelos propostos, será considerado aquele com menor valor de DIC, sendo simultaneamente o de melhor ajuste aos dados e melhor predição dos dados replicados, onde estes possuem a mesma estrutura dos dados observados.

Tabela 3.8 – Valores do DIC para comparação dos modelos de graduação propostos, segundo o ano de estudo, para as taxas unissex.

Ano 2004				
Modelo	$\bar{D}$	$\hat{D}$	$\rho D$	DIC
1	818,849	796,114	22,735	841,584
2	179,317	176,089	3,228	182,546
3	247,795	247,634	0,161	247,957
4	248,828	254,945	-6,117	242,710
Ano 2005				
Modelo	$\bar{D}$	$\hat{D}$	$\rho D$	DIC
1	799,628	776,757	22,871	822,499
2	158,868	155,383	3,485	162,353
3	226,428	226,373	0,055	226,482
4	228,337	232,740	-4,403	223,934
Ano 2006				
Modelo	$\bar{D}$	$\hat{D}$	$\rho D$	DIC
1	801,087	780,825	20,262	821,350
2	153,036	150,045	2,991	156,027
3	226,044	226,680	-0,636	225,408
4	227,973	233,443	-5,470	222,503
Ano 2007				
Modelo	$\bar{D}$	$\hat{D}$	$\rho D$	DIC
1	822,633	799,557	23,076	845,709
2	162,869	159,676	3,194	166,063
3	221,495	221,488	0,007	221,501
4	222,171	233,371	-11,200	210,971
Ano 2008				
Modelo	$\bar{D}$	$\hat{D}$	$\rho D$	DIC
1	780,203	763,712	16,491	796,693
2	179,489	175,088	4,402	183,891
3	227,781	227,877	-0,096	227,686
4	230,145	233,187	-3,042	227,102

Um dos requisitos para o uso do DIC é a convergência dos parâmetros de interesse através do modelo. Neste caso o requisito foi atendido.

Analisando os valores do DIC para cada modelo ajustado e ano de estudo, percebe-se que dentre os modelos locais, o modelo 2 é o que apresentou o menor valor do DIC em todos os anos, sendo preferível ao modelo 1. Todavia, os modelos globais 3 e 4 apresentaram valores do DIC inferiores aos modelos locais 1 e 2. Porém, no modelo 4 os valores da estatística  $pD$  são negativos em todos os anos de estudo, enquanto que para o modelo 3 os valores desta estatística são negativos apenas para os anos de estudo 2006 e 2008. Estes resultados negativos indicam bimodalidade da distribuição *a posteriori* do parâmetro, uma vez que os modelos globais se baseiam em MLG, em que a função de verossimilhança é sempre log-concava. Logo, de acordo com o DIC, concluímos que dentre os modelos locais, o modelo 2 é o que melhor se adéqua aos dados, ao passo que dentre os modelos globais, o modelo 3 é o preferido. Além disso, os valores da estatística  $\hat{R}$  para estes modelos indicaram convergência dos parâmetros de interesse em todas as idades (Figura 3.12), o que é um requisito para o uso do DIC na avaliação do modelo.

Por fim, a tábua de mortalidade bayesiana será construída considerando os modelos 2 e 3, dentre os quais será escolhido aquele que apresentar melhor adequabilidade aos dados segundo o método DIC.

### 3.4.3 CONSTRUÇÃO DA TÁBUA DE MORTALIDADE BAYESIANA

A construção de uma tábua de mortalidade deve levar em conta um período de estudo superior a um ano. Alguns autores utilizam em torno de quatro anos, sendo que, na elaboração de algumas tábuas bastante conhecidas, considerou-se um período de até seis anos, como por exemplo na AT-1949 e AT-1983. Também, as tábuas geralmente são separadas por sexo.

Para a construção da tábua de mortalidade da população em estudo, serão utilizados o modelo Bayesiano Estático Local e o modelo Bayesiano Estático Global, sendo o primeiro baseado no modelo 2 e o segundo baseado no modelo 3. O período considerado na construção da tábua de mortalidade compreende os anos de estudo de 2004 a 2008. Estes modelos não contemplam a incorporação da evolução das forças de mortalidade (parâmetros  $\theta_i$  a serem estimados) ao longo do tempo, nas classes de idade  $i$ , sendo os dados considerados com o mesmo peso em cada ano, caracterizando uma regressão com replicações em cada faixa etária. Quanto à nomenclatura “local” e “global”, se deve ao fato dos modelos 2 e 3 serem não paramétrico e paramétrico, respectivamente.



A estrutura dos dados, é  $D_t = (i, e_{i,t}, d_{i,t})$ , em que  $i$  é classe de idade,  $e_{i,t}$  é a quantidade central de expostos ao risco para a idade  $x+i-1$  no ano calendário  $t$  e  $d_{i,t}$  é o número de mortes observadas na idade  $x+i-1$  no ano calendário  $t$ , com classes de idade  $i=1, \dots, 49$ ,  $x=42$  e  $t=1, \dots, 5$ .

### 3.4.3.1 MODELO ESTÁTICO LOCAL

Como este modelo se baseia no modelo 2, aplica-se a mesma restrição à força de mortalidade, isto é,  $\theta \in R_I = \{\theta : 0 < \theta_1 < \theta_2 < \dots < \theta_{49} < 1\}$ , com a finalidade de graduar as probabilidades de morte estimadas. A função de verossimilhança dada em (3.2) é reescrita considerando os cinco anos de estudo, tomando a seguinte forma:

$$l(\theta | D_t) \propto \prod_{t=1}^T \prod_{i=1}^{49} \theta_i^{d_{i,t}} \exp(-e_{i,t} \cdot \theta_i) \quad (3.4)$$

em que  $\theta_i$  é a força de mortalidade para a classe de idade  $i$ ,  $T$  é o número de períodos de observação ( $T=5$ ),  $e_{i,t}$  e  $d_{i,t}$  já foram definidos anteriormente, com  $i=1, \dots, 49$  e  $t=1, \dots, T$ .

A distribuição a priori de  $\phi_i$  permanece a mesma atribuída ao modelo 2 e, uma vez especificadas a verossimilhança e a *priori*, tem-se a seguinte distribuição conjunta:

$$p(D_t, \phi | a, b) \propto \prod_{t=1}^5 \prod_{i=1}^{49} \theta_i^{d_{i,t}} \exp(-e_{i,t} \cdot \theta_i) \prod_{i=1}^{49} \phi_i^{a-1} \exp(-b \cdot \phi_i) \propto \prod_{t=1}^5 \prod_{i=1}^{49} \theta_i^{d_{i,t}} \phi_i^{a-1} \exp(-e_{i,t} \cdot \theta_i - b \phi_i)$$

Assim, a estrutura hierárquica bayesiana do Modelo Estático Local, é dada por:

$$\begin{aligned} d_{i,t} | \theta_i &\sim \text{Poisson}(e_{i,t} \cdot \theta_i), \text{ com } e_{i,t} \text{ conhecidos, } i = 1, \dots, 49 \text{ e } t = 1, \dots, T \quad (T = 5); \\ \theta_i | a, b &\sim \text{Gama}(a, b), \text{ com } a = 0,7830593 \text{ e } b = 231,27433; \\ \theta_i &= \theta_{i-1} + \phi_i \text{ e } \theta_0 = 0 \end{aligned}$$

com predição dada através de:

$$d_{i,t} | \theta_i \sim \text{Poisson}(e_{i,t} \cdot \theta_i), \text{ com } \theta_{i,T+1} = \frac{\sum_{t=1}^5 d_{i,t}^r}{\sum_{t=1}^5 e_{i,t}}$$

em que  $q_{x+i-1,T+1} = 1 - \exp(-\theta_{i,T+1})$ , para  $x = 42$

Para a implementação deste modelo no WinBUGS, por MCMC através do Amostrador de Gibbs, foram simuladas três cadeias em paralelo com valores iniciais dos  $\phi_i$ 's distintos, extraídos das tábuas de mortalidade AT-2000, AT-1983 e AT-1949, aplicando-se a relação (3.3). Adotou-se o tamanho das amostras geradas para cada parâmetro desconhecido como sendo 40.000, aproveitando uma a cada cinco simulações (*thinning*) e considerando as primeiras 20.000 como *burn-in*. Assim, o número de iterações considerado de fato, é de 8.000, das quais as primeiras 4.000 são desprezadas (*burn-in*), sendo as inferências a *posteriori* feitas sobre as últimas 4.000. O tempo gasto nas simulações das cadeias no WinBUGS, utilizando o computador já mencionado, foi de aproximadamente 60 minutos.

Para o diagnóstico da convergência dos parâmetros de interesse, foram utilizadas as mesmas técnicas informais de análise gráfica empregadas para os modelos anteriores.

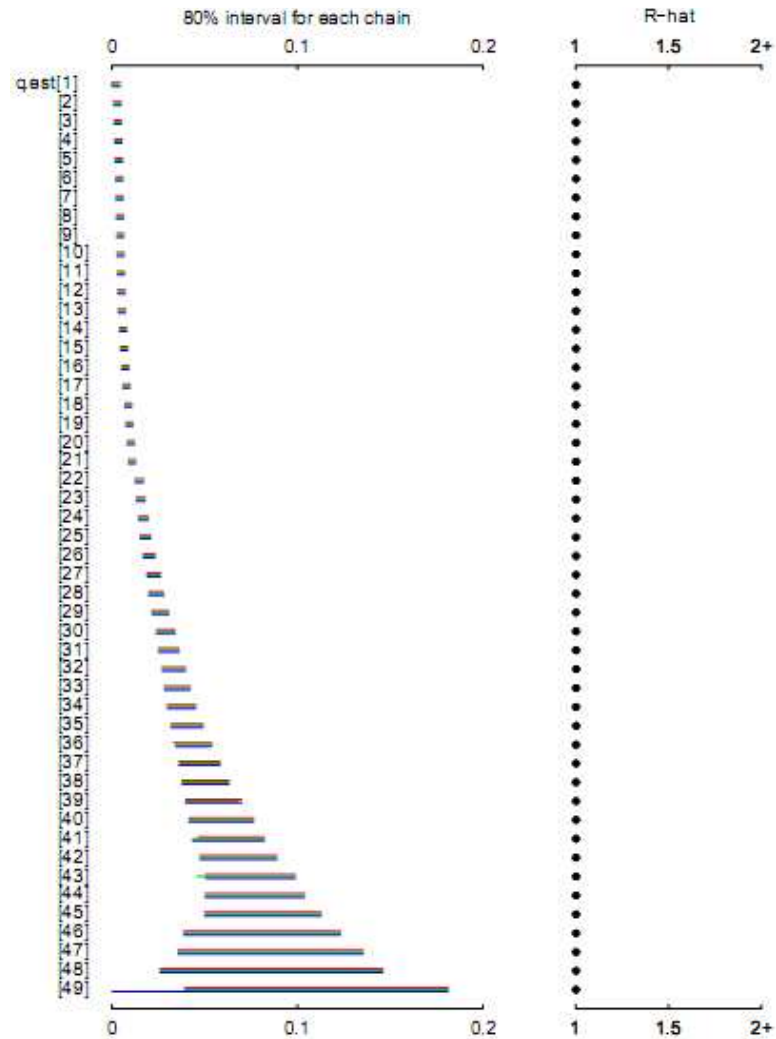


Figura 3.20 – Intervalos de credibilidade empíricos de 80% para as cadeias simuladas paralelamente e valores da estatística  $\hat{R}$  para análise da convergência dos parâmetros, para a tábua de mortalidade bayesiana.

Os valores das estatísticas  $\hat{R}$  (iguais a 1 para todas as estimativas  $\hat{q}_x$ ) na Figura 3.20 evidenciam as convergências dos parâmetros de interesse, isto é, a convergência das cadeias executadas em paralelo para os valores das probabilidades estimadas. Também, de forma semelhante ao modelo 2, os intervalos de credibilidade empíricos de 80% para cada parâmetro  $q_x$ , são bem estreitos nas idades mais jovens, se concentrando em torno de valores bem baixos de  $q_x$ , se tornando mais amplos à medida que a idade avança, devido aos mesmos motivos mencionados para o modelo 2. Aqui os intervalos de confiança são mais estreitos para as idades mais avançadas, acima de 83 anos.

Como já esperado, o comportamento das probabilidades estimadas ( $\hat{q}_x$ ) pelo Modelo Estático Local, continua suavizado e monotônico crescente. A Figura 3.21 permite uma melhor visualização deste comportamento.

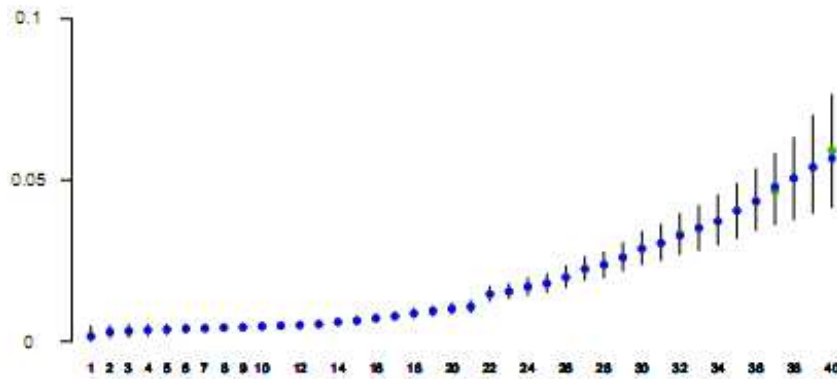


Figura 3.21 – Intervalos de credibilidade empíricos de 80% para as taxas de mortalidade da tábua Bayesiana, considerando as três cadeias simuladas em conjunto através do Modelo Estático Local.

Apesar do comportamento monotônico crescente esperado, nota-se da Figura 3.21 que o crescimento da taxa de mortalidade de 62 para 63 é um tanto brusco, de forma que entre estas idades o crescimento da taxa de mortalidade não é tão suave. As Figuras 3.22 e 3.23, apresentam respectivamente, um comparativo das probabilidades de morte da tábua bayesiana construída com aquelas usualmente adotadas nas EFPC's e o intervalo de confiança de 95% para a mesma.

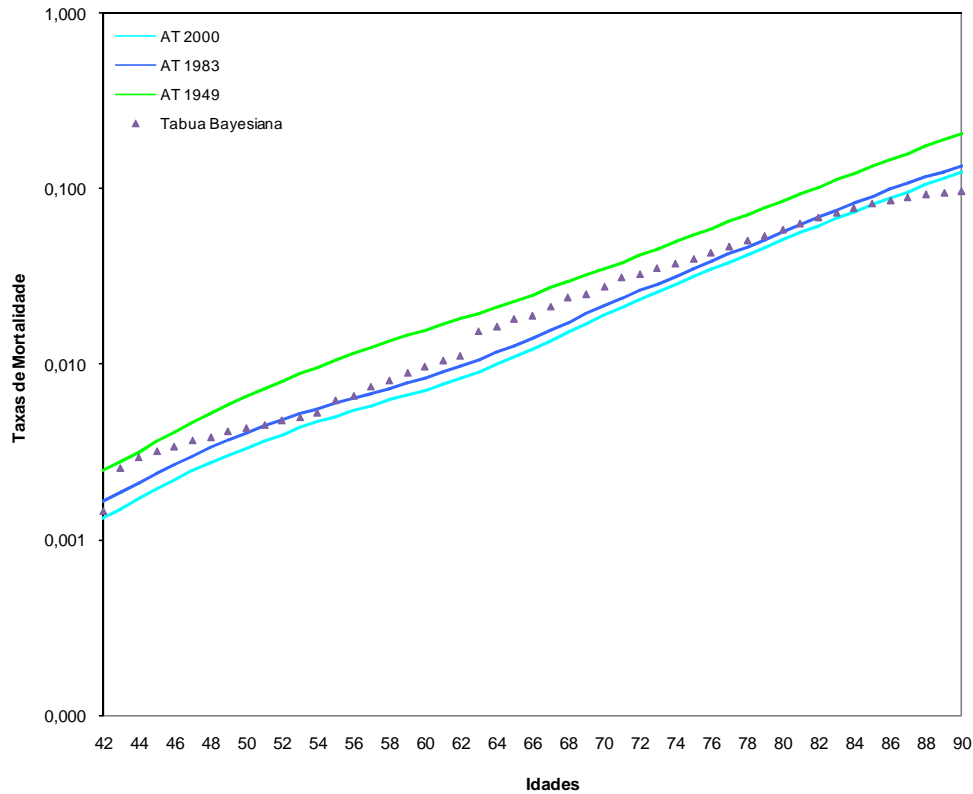


Figura 3.22 – Comparativo da tábua de mortalidade Bayesiana construída pelo Modelo Estático Local, considerando todos os anos de estudo, com as tábuas AT-2000, AT-1983 e AT-1949, em escala logarítmica.

Apesar do comportamento monotônico crescente apresentado pelas taxas de mortalidade graduadas, nota-se da Figura 3.22 que as mesmas oscilam bastante entre as tábuas AT-1949 e AT-1983. As taxas graduadas se comportam bem próximas às taxas de mortalidade da tábua AT-1949 nas idades mais jovens, se aproximando bastante da tábua AT-1983 entre aproximadamente 48 a 62 anos de idade, apresentando um crescimento brusco de 62 para 63 anos de idade, se tornando bem mais conservadoras a partir de aproximadamente 83 anos de idade e chegando a atingir patamares inferiores à tábua AT-2000. Embora a tábua bayesiana se torne bastante conservadora a partir de 85 anos, convém destacar que nestas idades existem muito poucos expostos ao risco, diminuindo a precisão das probabilidades de morte estimadas, conforme evidenciam os intervalos de credibilidade empíricos de 80% na Figura 3.20.

Enfim, as taxas graduadas pelo Modelo Estático Local atendem ao requisito da monotonicidade, porém no que diz respeito à suavidade, há uma nítida oscilação de seus valores, além de um crescimento brusco da taxa de mortalidade de 62 para 63 anos de idade, o qual é melhor visualizado em escala logarítmica na Figura 3.22.

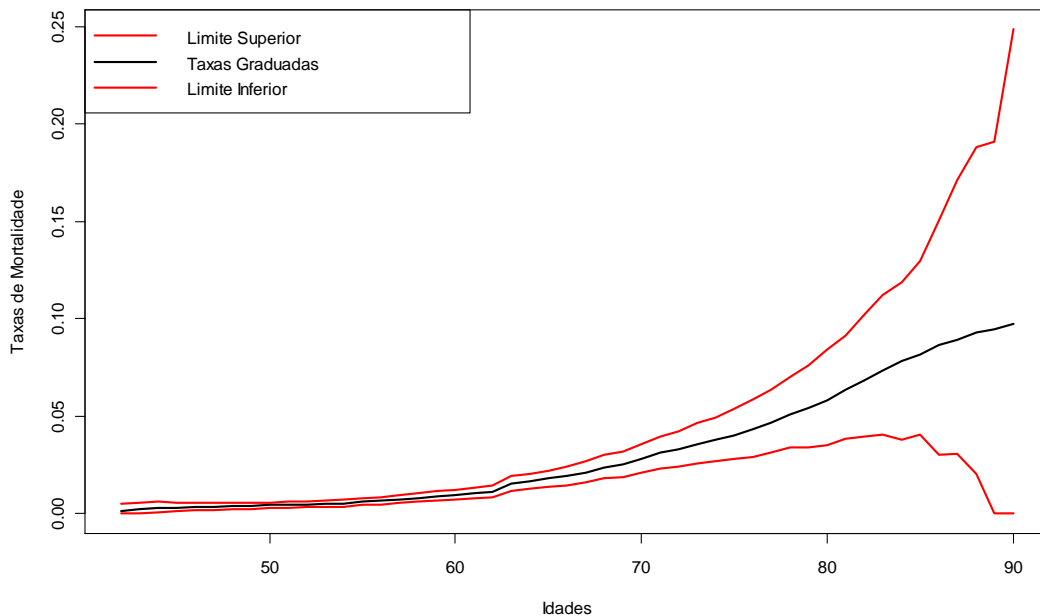


Figura 3.23 – Limites inferior e superior do intervalo de confiança de 95%, baseado em quantis, para a tábua de mortalidade bayesiana construída através do Modelo Estático Local.

O comportamento do intervalo de confiança de 95% para as probabilidades de morte da tábua bayesiana, extraído do WinBUGS, é semelhante aos intervalos de credibilidade empíricos de 80%, conforme a Figura 3.20: começam bem estreitos nas idades mais jovens se tornando gradativamente mais amplos conforme a idade avança, uma vez que há o aumento dos erros padrões das probabilidades de morte estimadas com este avanço. A partir de aproximadamente 85 anos, os intervalos se tornam extremamente amplos, indicando aumento acentuado da incerteza devido ao número bastante reduzido de expostos ao risco nestas idades.

Da mesma forma que na seção 3.4.2, também avaliamos as características do Modelo Estático Local segundo o DIC, o qual apresentou os seguintes valores:  $\bar{D} = 1144,78$ ,  $\hat{D} = 1132,5$ ,  $pD = 12,277$  e  $DIC = 1157,06$ . O valor da estatística  $\bar{D}$  reflete o grau de ajustamento do modelo aos dados, ao passo que a estatística  $pD$  - a qual é positiva - reflete o número efetivo de parâmetros do modelo e indica não haver conflito entre a distribuição *a priori* estabelecida e os dados, evidenciando também que a função de verossimilhança é log-côncava. Finalmente, o requisito de convergência para o uso deste método foi atendido, conforme os valores da estatística  $\hat{R}$  na Figura 3.20.

### 3.4.3.2 MODELO ESTÁTICO GLOBAL

Este modelo, da mesma forma que o modelo 3, pressupõe que a força de mortalidade obedece a Lei de Gompertz, isto é,  $\theta_i = \beta c^{z_i}$ , com  $Z_i = x+i-1$  representando as idades em estudo, onde  $x=42$  e  $i$  representa a  $i$ -ésima classe de idade, com  $i=1, \dots, 49$ . A função de verossimilhança para este modelo é dada por (3.4).

Como é assumido que o número de mortes na idade  $i$  e ano calendário  $t$ ,  $(d_{i,t} | \theta_i)$ , é distribuído segundo uma Poisson( $e_{i,t} \cdot \theta_i$ ), bem como o fato deste modelo estar alicerçado sobre MLG; a função de ligação canônica da distribuição de Poisson é dada por  $\eta_{i,t} = \ln(e_{i,t} \cdot \theta_i)$ . Logo, a estrutura do MLG é dada por:

$$\eta_{i,t} = \log(e_{i,t}) + \log(\beta) + Z_i \log(c)$$

em que  $e_{i,t}$  já foi definido na seção 3.4.3,  $t=1, \dots, T$ , sendo  $T$  o número de períodos de observação ( $T=5$ ).

Da mesma forma que no modelo 3, para possibilitar o processo de inferência bayesiana, atribuiu-se distribuições *a priori* Normais não informativas aos parâmetros  $\beta$  e  $c$ , ressaltando que no processo de simulação no pacote WinBUGS, utiliza-se o parâmetro de precisão  $\tau = \sigma^{-2}$  ao invés da variância. Assim, a estrutura hierárquica bayesiana do Modelo Estático Global, é dada por:

$$\begin{aligned} d_{i,t} | \theta_i &\sim \text{Poisson}(\lambda_{i,t}), \lambda_{i,t} = e_{i,t} \cdot \theta_i, \text{ com } e_{i,t} \text{ conhecidos e } i = 1, \dots, 49, t = 1, \dots, 5; \\ \log(\lambda_{i,t}) &= \log(e_{i,t}) + \log(\beta) + Z_i \log(c); \\ \theta_i &= \beta c^{Z_i} \\ \beta &\sim \text{Normal}(0; 10000) I_{(0, \infty)}(\beta); \\ c &\sim \text{Normal}(0; 10000) I_{(1, \infty)}(c) \end{aligned}$$

com predição dada por:

$$d_{i,t} | \theta_i \sim \text{Poisson}(e_{i,t} \cdot \theta_i), \text{ com } \theta_{i,T+1} = \frac{\sum_{t=1}^5 d_{i,t}^r}{\sum_{t=1}^5 e_{i,t}}$$

em que  $q_{x+i-1, T+1} = 1 - \exp(-\theta_{i,T+1})$ , para  $x = 42$

Para a implementação deste modelo no WinBUGS, por MCMC, através do Amostrador de Gibbs, foi simulada apenas uma cadeia. O tamanho das amostras geradas para cada parâmetro desconhecido foi de 40.000, aproveitando uma a cada cinco simulações (*thinning*) e considerando as primeiras 20.000 como *burn-in*. Logo, o número de iterações aproveitadas foi de 8.000, das quais as primeiras 4.000 foram desprezadas (*burn-in*), sendo as inferências a *posteriori* feitas sobre as últimas 4.000. O tempo gasto na simulação, utilizando o computador já mencionado, foi de aproximadamente 2 minutos, sendo o diagnóstico da convergência dos parâmetros de interesse realizados através das mesmas técnicas informais de análise gráfica empregadas nos modelos anteriores.

O comportamento das probabilidades estimadas ( $\hat{q}_x$ ) pelo Modelo Estático Global, é dado na Figura 3.24.

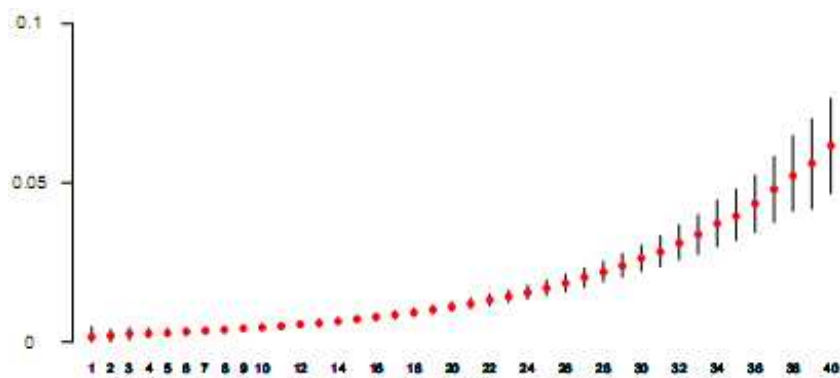


Figura 3.24 – Intervalos de credibilidade empíricos de 80% para as taxas de mortalidade da tábua Bayesiana, considerando uma cadeia simulada através do Modelo Estático Global.

As taxas graduadas através do Modelo Estático Global apresentaram comportamento dentro do esperado, isto é, monotônico crescente e suavizado, ao contrário do Modelo Estático Local, que apesar de apresentar a monotonicidade esperada, não proporciona a suavidade desejada no crescimento das taxas de mortalidade, conforme mostra a Figura 3.21. Também, quando comparado o Modelo Estático Global com o modelo 3, este apresenta menores amplitudes dos intervalos de credibilidade, diminuindo portanto os erros padrões das estimativas das probabilidades de morte estimadas ( $\hat{q}_x$ ), aumentando a precisão das estimativas. As Figuras 3.25 e 3.26 apresentam respectivamente, um comparativo das probabilidades de morte da tábua bayesiana construída com aquelas usualmente adotadas nas EFPC's e o intervalo de confiança de 95% para a mesma.



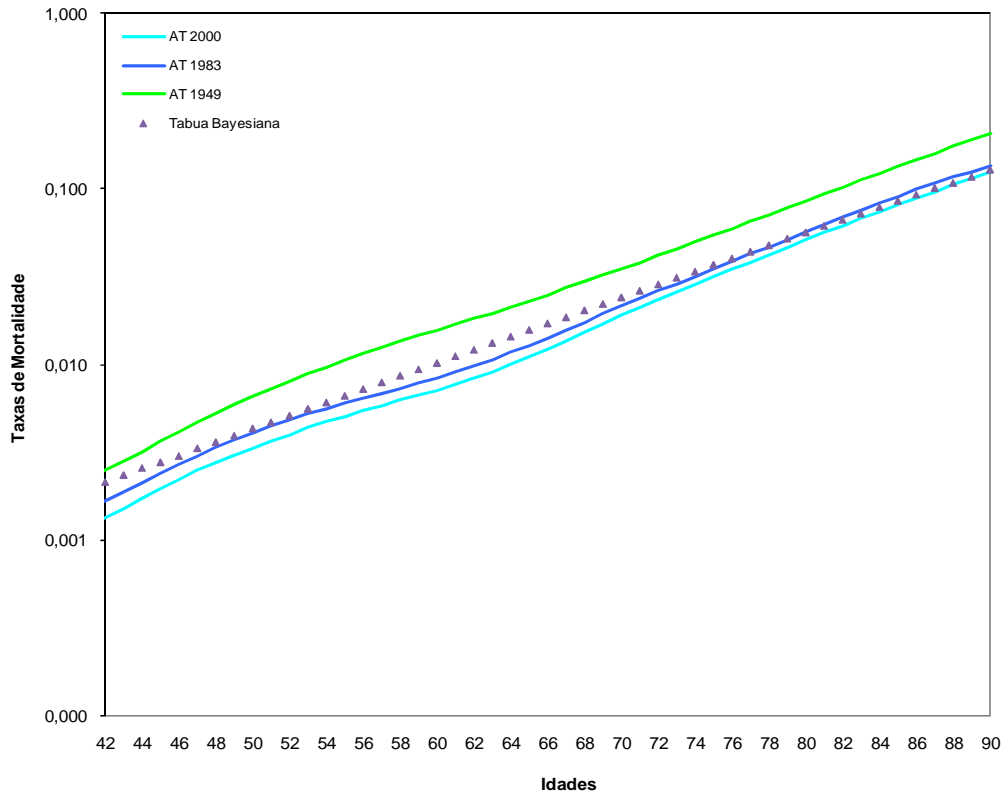


Figura 3.25 – Comparativo da tábua de mortalidade Bayesiana construída pelo Modelo Estático Global, considerando todos os anos de estudo, com as tábuas AT-2000, AT-1983 e AT-1949, em escala logarítmica.

Nota-se da Figura 3.25 que comportamento apresentado pelas taxas graduadas através do Modelo Estático Global, além de apresentar o comportamento monotônico crescente esperado, também atende ao quesito da suavidade no crescimento das taxas, não apresentando qualquer crescimento brusco nem oscilações como aquelas ocorridas na graduação realizada através do Modelo Estático Local. As taxas graduadas por este modelo permanecem bem próximas daquelas provenientes da tábua AT-1983, vindo de encontro ao resultado do teste de aderência realizado para esta tábua na seção 3.3, se tornando gradativamente mais conservadoras entre aproximadamente 80 e 90 anos de idade, quando se aproximam bastante da tábua AT-2000. Portanto, com base nos diagnósticos gráficos, as taxas graduadas pelo Modelo Estático Global apresentam comportamento mais coerente que aquelas graduadas pelo Modelo Estático Local.

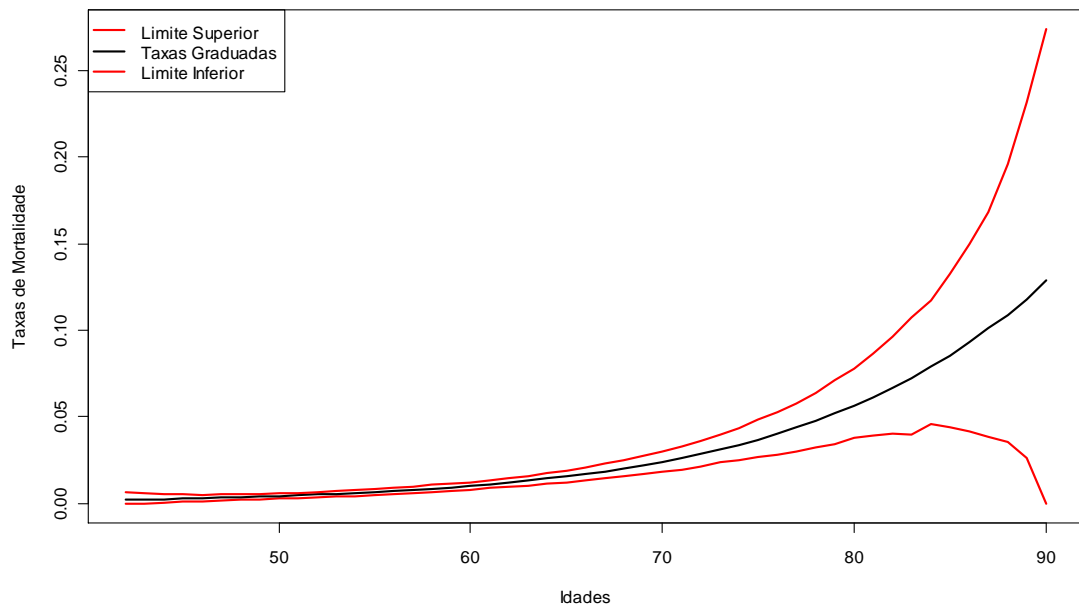


Figura 3.26 – Limites inferior e superior do intervalo de confiança de 95%, baseado em quantis, para a tábua de mortalidade bayesiana construída através do Modelo Estático Global.

O comportamento do intervalo de confiança de 95% para as probabilidades de morte da tábua bayesiana construída através do Modelo Estático Global, é semelhante aos intervalos de credibilidade empíricos de 80%, conforme a Figura 3.24: começam bem estreitos nas idades mais jovens se tornando gradativamente mais amplos conforme a idade avança, uma vez que há o aumento dos erros padrões das probabilidades de morte estimadas com este avanço. Da mesma forma que para o Modelo Estático Local, a partir de aproximadamente 85 anos, os intervalos se tornam extremamente amplos, indicando diminuição da precisão das estimativas das probabilidades de morte, devido ao número bastante reduzido de expostos ao risco nestas idades. Apesar desta semelhança com o Modelo Estático Local, aqui o intervalo de confiança é um pouco mais estreito, além de apresentar limites de confiança mais suavizados.

Segundo o método *DIC*, foram obtidos os seguintes valores para este modelo:  $\bar{D} = 1164,62$ ,  $\hat{D} = 1162,96$ ,  $pD = 1,659$  e  $DIC = 1166,28$ . Como se pode observar, com exceção da estatística *pD*, os valores são um pouco superiores àqueles obtidos para o Modelo Estático Local. Portanto, apesar do valor do *DIC* para este modelo ter sido um pouco superior, a diferença não é significativa, de forma que este modelo foi o escolhido. Além disso, os diagnósticos gráficos são mais favoráveis a adoção deste modelo.

#### 3.4.4 ANÁLISE DOS RESULTADOS

O comportamento apresentado pela tábua bayesiana construída pelo Modelo Estático Global mostrou-se bastante próximo à tábua de mortalidade AT-1983, se tornando gradativamente mais conservadora nas idades acima de 80 anos. Este comportamento também apresenta coerência com o resultado do teste de aderência realizado para a tábua AT-1983, a qual se mostrou aderente à população exposta ao risco, reforçando a determinação da Resolução CGPC N° 18, a qual adota esta tábua como parâmetro mínimo de mortalidade geral nas EFPC's. Também, a tábua construída possibilita sua aplicação em cálculos atuariais, uma vez que apresentou o crescimento monotônico e suavidade esperada das taxas graduadas. Todavia, para as idades a partir de 85 anos, a falta de precisão destas estimativas é bastante grande, devido ao número bastante reduzido de expostos ao risco nestas idades. Logo, devido a este fato, não seria recomendada a aplicação da tábua construída para as idades entre 85 e 90 anos.

A despeito da tábua de mortalidade bayesiana ter sido construída com base na experiência da população exposta ao risco, pelo princípio da tendência de quedas nas taxas de mortalidade da população mundial - já comprovado através de diversos estudos - os indicativos são de adoção de hipóteses de mortalidade cada vez mais conservadoras, isto é, menores taxas de mortalidade. Devido a esta evolução dos níveis de mortalidade, a adoção de tábuas geracionais vem se tornando popular nas entidades de previdência e seguridade. Desta forma, no próximo capítulo avaliaremos a aplicação do modelo de Lee-Carter na projeção das taxas de mortalidade da tábua bayesiana construída, tornando-a na forma geracional.

#### 4 PROJEÇÃO DE TAXAS DE MORTALIDADE

Um dos principais desafios impostos pelo fenômeno do envelhecimento populacional é manter a aderência das hipóteses de mortalidade adotadas num sistema de previdência e seguridade. Dado que a tábua de mortalidade é uma ferramenta imprescindível no cálculo atuarial, a aderência aos níveis de mortalidade de uma determinada população é um requisito fundamental. Todavia, a tábua de mortalidade é estática e, devido aos movimentos dos níveis de mortalidade que vêm sendo observados em decorrência do rápido envelhecimento populacional; poderá ser descartada num prazo relativamente curto. Uma das formas de minimizar, ou mesmo neutralizar, os impactos decorrentes destes movimentos, é a obtenção de informações sobre o comportamento da mortalidade futura através de métodos de projeção, de forma a contemplar essas tendências no cálculo atuarial.

De acordo com Santos (2007), os métodos de projeção extrapolativos são os mais adequados, se classificando em:

- paramétricos: envolvem o ajuste de uma curva/superfície parametrizada aos dados para os anos, projetando estes parâmetros para o futuro;
- *targeting*: envolvem interpolação entre as taxas de mortalidade atuais e um conjunto de taxas esperadas, as quais se assume que vão permanecer em uma data futura;
- tendência: envolvem a projeção de tendências históricas para o futuro.

O modelo de Lee-Carter, detalhado na seção 2.3 é um método extrapolativo baseado em projeção de tendências históricas e utilizado neste trabalho para obter os fatores de melhoria (*improvement*) a serem aplicados sobre as taxas de mortalidade de tábuas estáticas, projetando-as para o futuro. Uma componente deste modelo ( $k_t$ ) é tipicamente ajustada por um modelo de séries temporais da classe ARIMA (BOX; JENKINS, 1976), a qual tem como pressuposto, um histórico suficientemente grande dos dados em anos de observação, uma vez que esta classe de modelos é baseada em resultados assintóticos.

No Brasil, é possível a obtenção de dados censitários publicados por órgãos de pesquisa do governo. Porém é sabido que estas estatísticas são uma mesclagem de perfis demográficos diferentes, não sendo uma boa base para o comportamento de populações específicas, tais como das EFPC's (SANTOS; 2007). Portanto, em se tratando de populações

das EFPC's, ainda não há disponibilidade de um histórico de dados suficientemente longo para um bom ajuste da série.

Devido à ausência de um histórico de dados adequado, nos baseamos no estudo de Santos (2007), que obteve os fatores de melhoria aplicando o modelo de Lee-Carter aos dados da Inglaterra e País de Gales, considerando o período de 1950 a 2003 e as idades de 18 a 100 anos. Neste trabalho utilizamos os mesmos dados, porém para as idades de 42 a 90 anos, conforme as idades da população exposta ao risco considerada neste estudo. Também, uma vez ajustado o modelo ARIMA à componente  $k_t$  do modelo de Lee-Carter, realizamos simulações estocásticas do mesmo através de *bootstrap* paramétrico, construindo intervalos de confiança empíricos baseados em 99 simulações para as taxas centrais de mortalidade geradas, a partir dos  $k_t$  simulados. Assim, verificamos a possibilidade de aplicação dos fatores de melhoria obtidos a partir dos  $k_t$  projetados pelo modelo ARIMA ajustado aos dados da Inglaterra e País de Gales, na população considerada neste estudo.

#### 4.1 TÁBUAS ESTÁTICAS E GERACIONAIS

Como o próprio nome diz, uma tábua estática não contempla os movimentos ocorridos ao longo do tempo nos níveis de mortalidade. Desta forma, se considerarmos uma determinada idade  $x$ , a probabilidade de morte e expectativa de vida nesta idade, num determinado ano calendário  $t$ , permanece a mesma no ano calendário  $t+k$  futuro, ignorando qualquer movimento nos níveis de mortalidade. Em contrapartida, uma tábua geracional, nada mais é do que uma tábua estática sobre a qual se aplicam escalas de projeção (fatores de melhoria) em suas respectivas probabilidades, de forma a contemplar a evolução dos níveis de mortalidade ao longo do tempo.

Uma tábua de mortalidade geracional que tem sido utilizada em algumas EFPC's, é a RP-2000, a qual foi publicada pela SOA (*Society of Actuaries*). Tal tábua foi construída considerando o período de observação de 1990 a 1994, baseando-se na experiência de 113 Fundos de Pensão americanos. A escala de projeção a ser aplicada nas probabilidades de morte desta tábua, de forma a torná-la geracional, é denominada escala AA, a qual deve ser aplicada a partir do ano-base de 2000. Apesar de serem considerados 4 anos na construção da RP-2000, a escala AA foi construída com base num período de observação de 17 anos.

A Figura 4.1, extraída de Santos (2007), ilustra a evolução das probabilidades de morte, segundo uma tábua estática e segundo uma tábua geracional.

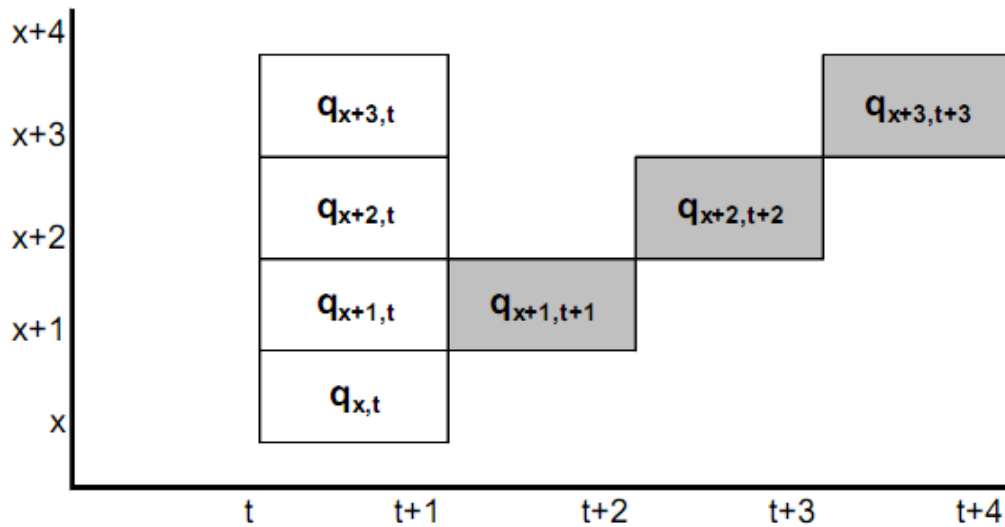


Figura 4.1 – Comparação das probabilidades de morte de uma tabua estática (quadro em branco) com as de uma tábua geracional (quadro cinza) (SANTOS, 2007).

O eixo horizontal da Figura 4.1 representa os anos calendário, enquanto o eixo vertical representa as idades. Percebe-se que à medida que os anos calendário aumentam, as probabilidades de morte ( $q_x$ ) também mudam (quadros cinza), decorrentes da aplicação dos fatores de melhoria  $FI(x,t)$  sobre as mesmas. Assim, tem-se:

$$q_{x,1} = FI(x,1) \cdot q_x$$

$$q_{x,2} = FI(x,2) \cdot q_x$$

...

$$q_{x,t} = FI(x,t) \cdot q_x$$

## 4.2 APLICAÇÃO DO MODELO DE LEE-CARTER

Assim como em Santos (2007), neste estudo o modelo de Lee-Carter também foi aplicado aos dados da população da Inglaterra e País de Gales, para o período de 1950 a 2003. Todavia, a modelagem foi unissex (não segregada por sexo), considerando a faixa etária entre 42 e 90 anos, de acordo com os dados da população das EFPC's componentes deste estudo.

Para o ajuste do modelo, conforme equação (2.9), realizou-se a estimação dos parâmetros  $a_x$ ,  $b_x$  e  $k_t$  numa primeira etapa, reestimando  $k_t$  posteriormente, de forma que o número total de óbitos ajustado coincida com o número total de óbitos efetivamente verificados. Os procedimentos da estimação são dados na próxima seção.

#### 4.2.1 ESTIMAÇÃO DOS PARÂMETROS DO MODELO

Primeiramente procedeu-se à estimação dos parâmetros  $a_x$ , através dos logaritmos das médias geométricas das taxas centrais de mortalidade em cada idade  $x$ , avaliadas sobre todo o ano calendário  $t$ , através da seguinte equação:

$$\hat{a}_x = \frac{1}{h} \sum_{t=t_1}^{t_n} \ln(m_{x,t}) = \ln \left[ \prod_{t=t_1}^{t_n} m_{x,t}^{1/h} \right] \quad (4.1)$$

em que  $h=t_n-t_1$  e  $m_{x,t}$  é a taxa central de mortalidade para a idade  $x$  no tempo  $t$ .

Uma vez obtidas as estimativas  $\hat{a}_x$ , procede-se a primeira etapa da estimação dos parâmetros  $k_t$ , de acordo com a seguinte equação:

$$\hat{k}_t = \sum_{x=48}^{90} [\ln(m_{x,t}) - \hat{a}_x] \quad (4.2)$$

Para a estimação dos parâmetros  $\hat{b}_x$ , é feita a regressão de  $[\ln(m_{x,t}) - \hat{a}_x]$  em  $\hat{k}_t$ , sem o termo constante, para cada idade  $x$ , através da seguinte equação:

$$\ln(m_{x,t}) - \hat{a}_x = b_x \hat{k}_t^1 + \varepsilon_{x,t} \quad (4.3)$$

em que o sobrescrito 1 em  $\hat{k}_t^1$  se refere à primeira etapa da estimação deste parâmetro e,  $\varepsilon_{x,t}$ , é um erro aleatório com distribuição Normal de média 0 e variância  $\hat{\sigma}_\varepsilon^2$ , conforme o modelo de regressão linear simples ajustado pelo método dos mínimos quadrados. Logo, normalizando a equação a equação (4.3), encontramos o estimador de mínimos quadrados de  $b_x$ , da seguinte forma:

$$\sum_{x,t} [\ln(m_{x,t}) - \hat{a}_x - b_x \hat{k}_t^1]^2 = \sum_{x,t} (\varepsilon_{x,t})^2 \quad (4.4)$$

Fixando  $x$ , derivando (4.4) em relação à  $b_x$  e igualando a 0, tem-se:

$$\frac{\partial}{\partial b_x} \sum_{x,t} [\ln(m_{x,t}) - \hat{a}_x - b_x \hat{k}_t^1]^2 = 0 \Rightarrow \hat{b}_x = \frac{\sum_{t=t_1}^{t_n} \hat{k}_t^1 [\ln(m_{x,t}) - \hat{a}_x]}{\sum_{t=t_1}^{t_n} (\hat{k}_t^1)^2} \quad (4.5)$$

A Figura 4.2 ilustra os comportamentos das estimativas  $\hat{a}_x$ ,  $\hat{k}_t^1$  e  $\hat{b}_x$ , obtidas através de (4.1), (4.2) e (4.5), respectivamente.

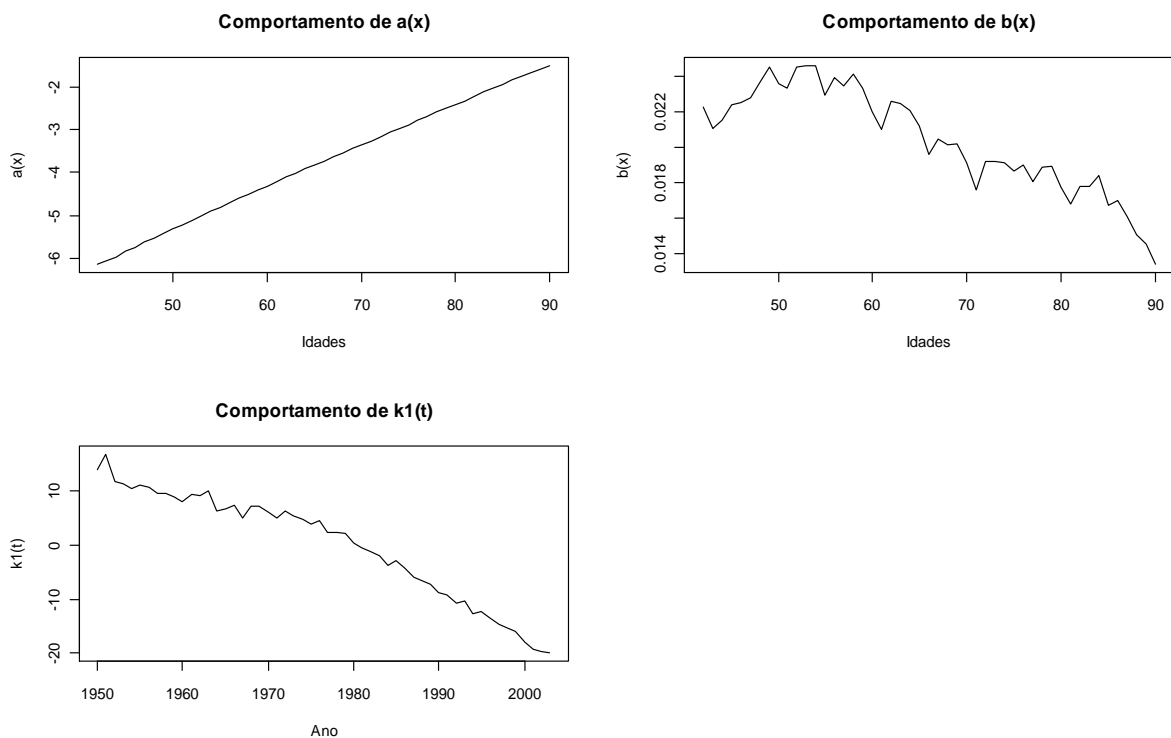


Figura 4.2 – Comportamentos das estimativas  $\hat{a}_x$  e  $\hat{b}_x$  para as idades de 42 a 90 anos e comportamento das estimativas  $\hat{k}_t^1$  para os anos calendário de 1950 a 2003, unissex.

O comportamento das estimativas  $\hat{a}_x$ , apresentam uma forte tendência linear, aumentando à medida que a idade avança. Quanto às estimativas  $\hat{k}_t^1$ , é constatada a tendência de queda no nível de mortalidade com o avanço no tempo. Finalmente, as estimativas  $\hat{b}_x$ , que avaliam a sensibilidade da variação nas taxas de mortalidade conforme o parâmetro  $k_t$  varia, apresentam uma tendência decrescente, de forma que as taxas de mortalidade variam pouco



quando o nível geral de mortalidade se altera (idades mais avançadas), ao passo que são bem mais sensíveis a variação do parâmetro  $k_t$  nas idades mais jovens.

Após a realização desta primeira etapa, os parâmetros  $k_t$  são reestimados, utilizando os valores das estimativas  $\hat{a}_x$  e  $\hat{b}_x$  já obtidos. Assim, as estimativas de  $k_t$  na segunda etapa, devem ser tal que a equação (2.10) seja atendida, a qual está reproduzida abaixo:

$$D_t = \sum_x [N_{x,t} \exp(\hat{a}_x - k_t b_x)]$$

Em que  $D_t$  é o número total de óbitos observados no ano  $t$  e  $N_{x,t}$  corresponde à população exposta ao risco com idade  $x$ , no mesmo ano  $t$ .

Conforme descrito em Santos (2007) há muitas vantagens em se fazer a estimação de segundo estágio do parâmetro  $k_t$ . A primeira razão é que isto garante que as taxas de mortalidade ajustadas sobre os anos amostrados vão se ajustar ao número total de mortes e a distribuição etária da população. Dado que na estimação de primeiro estágio de  $k_t$ , baseou-se no logaritmo das taxas centrais de mortalidade e não nas mesmas, existe a possibilidade de discrepâncias entre o número de mortes observado e o estimado. A segunda é que, em alguns períodos, a distribuição etária da população e o número total de mortes sejam conhecidos, mas as taxas de mortalidade específicas por idade não.

A Figura 4.3 ilustra o comportamento das estimativas de 1º e 2º estágios de  $k_t$ , dadas por  $\hat{k}_t^1$  e  $\hat{k}_t^2$ , respectivamente. Nota-se, que os comportamentos de  $\hat{k}_t^1$  e  $\hat{k}_t^2$  são muito semelhantes, de forma que os valores destas estimativas apresentam um declínio aproximadamente linear. Esta linearidade aproximada e a variação relativamente constante da série  $\hat{k}_t^2$  no período considerado, são consideradas vantajosas na construção de modelos de previsão. Os valores das estimativas  $\hat{a}_x$ ,  $\hat{b}_x$ ,  $\hat{k}_t^1$  e  $\hat{k}_t^2$  estão disponíveis no Anexo V.

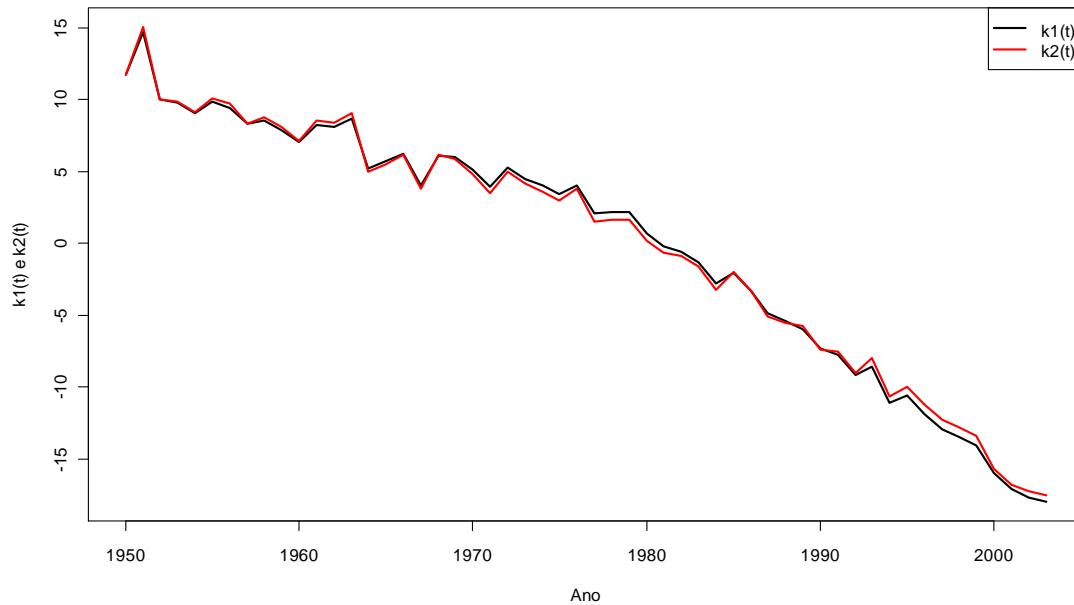


Figura 4.3 – Comparativo entre os comportamentos das estimativas  $\hat{k}_t^1$  e  $\hat{k}_t^2$ , considerando os anos de observação 1950 a 2003

Uma vez reestimado  $k_t$ , pode-se escrever o modelo demográfico proposto na equação (2.9) e, como os parâmetros  $a_x$  e  $b_x$  estimados não dependem da variável tempo, a projeção das taxas de mortalidade futuras dependerão exclusivamente dos valores futuros de  $k_t$ . Para se projetar estes valores,  $k_t$  pode ser modelado por um processo da classe ARIMA, conforme a metodologia Box & Jenkins (1976).

#### 4.2.2 AJUSTE DE UM PROCESSO ARIMA À COMPONENTE $k_t$ DO MODELO

Preliminarmente, um modelo ARIMA possui três parâmetros:  $p$ ,  $d$ , e  $q$ , de forma que a notação usual para este modelo é: ARIMA( $p,d,q$ ). O parâmetro  $p$  corresponde à ordem da parte autorregressiva do modelo - AR( $p$ ) - enquanto o parâmetro  $q$  corresponde à ordem da parte médias móveis do modelo - MA( $q$ ). Já o parâmetro  $d$ , corresponde à ordem de integração do modelo, isto é, indica quantas vezes a série é diferenciada, uma vez que este artifício as vezes é usado para estacionarizar séries não estacionárias.

Anteriormente ao ajuste do modelo, foram feitas algumas inspeções na série  $\hat{k}_t^2$ , a fim de identificar algumas características que poderão auxiliar na escolha do modelo adequado para a componente  $k_t$  do modelo de Lee-Carter.

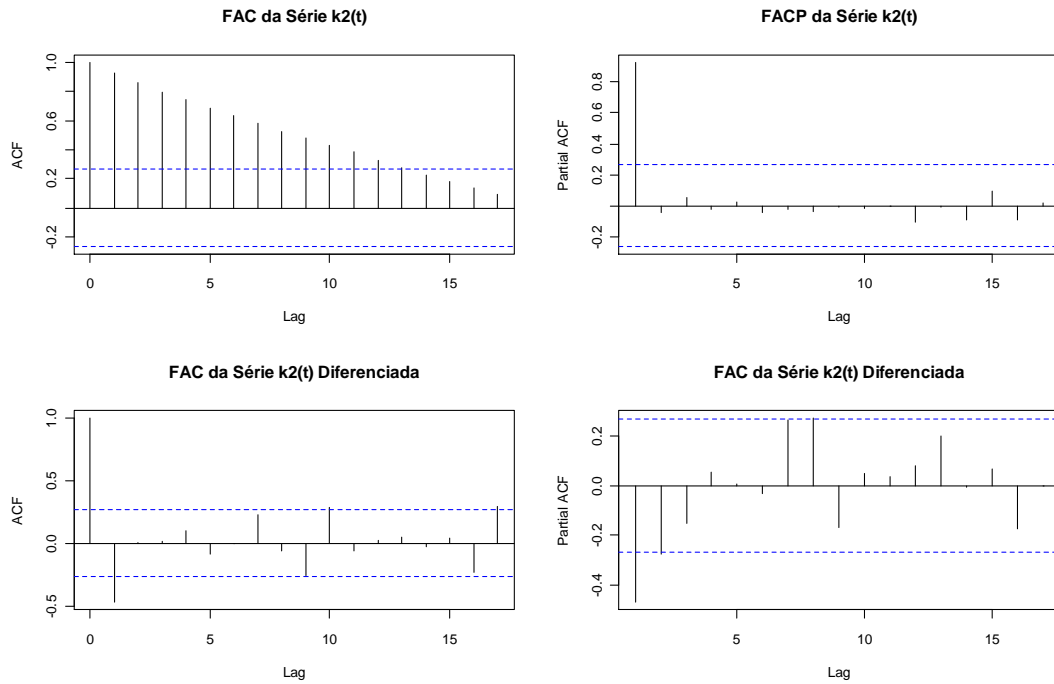


Figura 4.4 – Funções de autocorrelação (FAC) e autocorrelação parcial (FACP) das séries original e diferenciada das estimativas de segundo estágio  $\hat{k}_t^2$ .

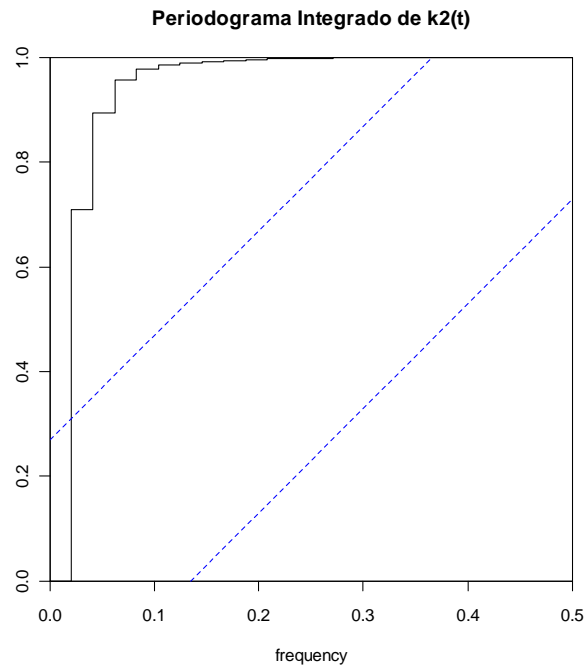


Figura 4.5 – Periodograma Integrado das estimativas de segundo estágio  $\hat{k}_t^2$ .

Como se observa da Figura 4.4, a função de autocorrelação (FAC) de  $\hat{k}_t^2$ , evidencia um comportamento altamente correlacionado das observações da série até o lag 13,

apresentando independência entre os valores da série somente a partir do lag 14. Isto é devido à tendência linear decrescente apresentada por  $\hat{k}_t^2$ . Portanto, a série é não estacionária no nível. Já a função de autocorrelação parcial (FACP), a qual auxilia na escolha do modelo, inicialmente sugere o ajuste de um modelo autorregressivo de ordem 1 – AR(1) – devido a presença de pico (*spike*) no lag 1. Todavia, no periodograma integrado da série  $\hat{k}_t^2$ , Figura 4.5, os valores das ordenadas (linha preta) extrapolam o limite de confiança superior (linha azul pontilhada), descartando a hipótese de ajuste da série  $\hat{k}_t^2$  a um modelo AR(1). Desta forma, a série foi diferenciada uma vez, eliminando-se a tendência decrescente apresentada pela série original. Logo, o modelo que melhor se adequou a série  $\hat{k}_t^2$  foi o ARIMA(0,1,1), isto é, modelo de médias móveis (MA) diferenciado uma vez. Portanto, o modelo ajustado é dado pela seguinte expressão:

$$\Delta k_t = \hat{\delta} + \hat{\theta}_1 \gamma_{t-1} + \gamma_t$$

em que o ajuste retornou as seguintes estimativas dos parâmetros: termo constante  $\hat{\delta} = -0,56664$ , parâmetro média móvel  $\hat{\theta}_1 = 0,551861$ , com  $\hat{\sigma}_\gamma^2 = 1,47595$  sendo a estimativa da variância dos erros aleatórios (resíduos)  $\gamma_t$ , que são normalmente distribuídos com média 0, representando uma série “ruído branco” completamente aleatória, confirmado através do periodograma na Figura 4.6.

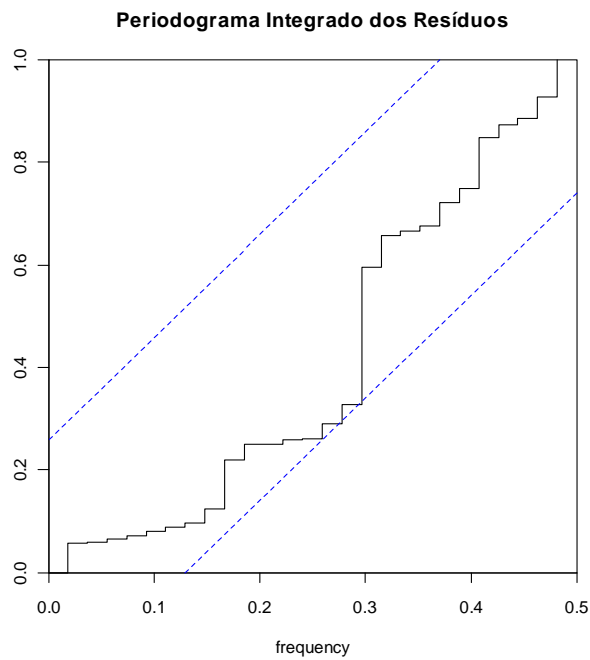


Figura 4.6 – Periodograma Integrado da série dos erros aleatórios  $\gamma_t$  disponibilizado pelo pacote R.

Os valores das ordenadas no períodograma integrado se comportam dentro dos limites de confiança, reforçando a evidência de que os  $\gamma_t$  apresentam comportamento completamente aleatório.

A Figura 4.7 evidencia a normalidade dos erros  $\gamma_t$  e os diagnósticos confirmam a adequabilidade do modelo ARIMA(0,1,1) ao ajuste da componente  $k_t$  do modelo de Lee-Carter.

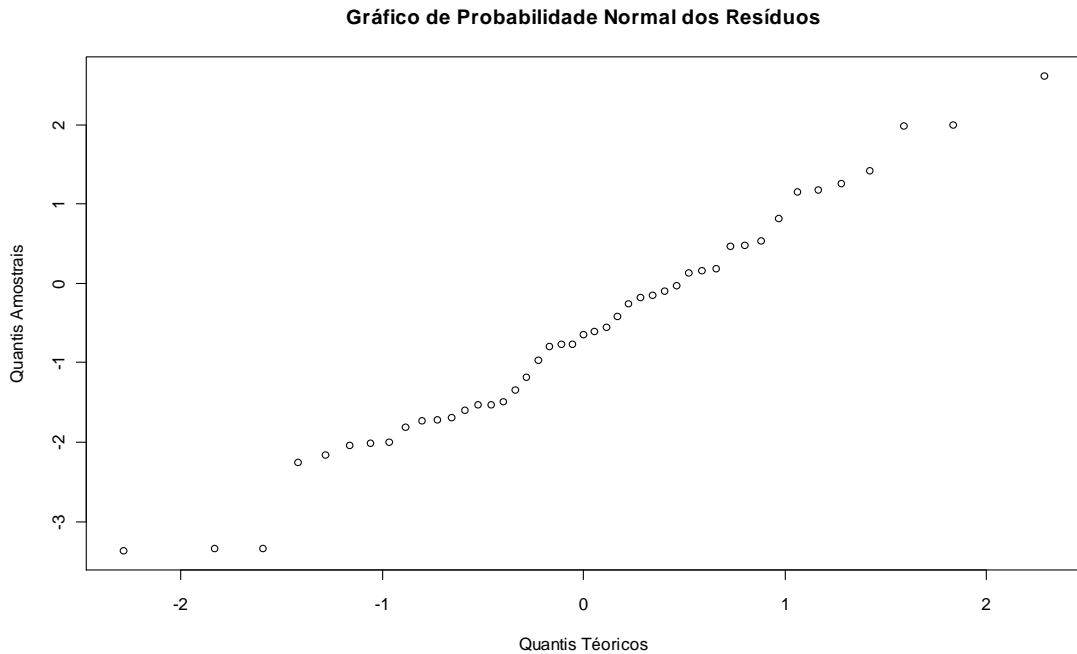


Figura 4.7 – Gráfico de probabilidade Normal dos erros aleatórios.

A Tabela 4.1 contém os valores projetados de  $k_t$  para os próximos 12 anos, a partir de 2003.

Tabela 4.1 – Valores projetados de  $k_t$  com os limites de predição de 95%, para os próximos 12 anos, disponibilizados pelo pacote R.

Ano (t)	$k_t$	Limites	
		Inferior	Superior
2004	-18,0013	-20,4403	-15,5623
2005	-18,5680	-21,2407	-15,8952
2006	-19,1346	-22,0222	-16,2470
2007	-19,7012	-22,7887	-16,6137
2008	-20,2679	-23,5431	-16,9926
2009	-20,8345	-24,2873	-17,3817
2010	-21,4012	-25,0228	-17,7795
2011	-21,9678	-25,7508	-18,1848
2012	-22,5344	-26,4722	-18,5967
2013	-23,1011	-27,1877	-19,0144
2014	-23,6677	-27,8980	-19,4374
2015	-24,2343	-28,6036	-19,8651

Uma vez modelado  $k_t$ , a partir de seus valores preditos pode-se também projetar as taxas centrais de mortalidade, de forma que as tábuas de mortalidade sejam construídas. De

acordo com Santos (2007) e segundo o modelo de Lee-Carter, a partir da última taxa central de mortalidade disponível para uma determinada idade  $x$ , isto é,  $m_{x,2003}$ , pode-se projetar a mesma para  $n$  períodos à frente, conforme a seguinte expressão:

$$\hat{m}_{x,2003+n} = \hat{m}_{x,2003} \cdot FI(x,2003+n) = \hat{m}_{x,2003} \cdot \exp\left[\hat{b}_x (\hat{k}_{2003+n} - \hat{k}_{2003})\right] \quad (4.6)$$

em que  $FI(x,2003+n)$  é o fator de *improvement* a ser aplicado sobre a taxa central de mortalidade e  $n$  é um inteiro maior ou igual a 0.

A partir das taxas centrais de mortalidade projetadas, obtêm-se as probabilidades de morte  $q_x$  que comporão a tábua de mortalidade, a partir da seguinte relação:

$$q_x = \frac{2m_x}{2+m_x} \Rightarrow q_{x,t} = \frac{2m_{x,t}}{2+m_{x,t}} \quad (4.7)$$

Neste trabalho as taxas de mortalidade foram graduadas anteriormente, quando da construção da tábua de mortalidade bayesiana, sendo que a projeção feita através do modelo de Lee-Carter é sobre as taxas centrais de mortalidade  $m_x$  ao invés das taxas graduadas  $q_x$ . Desta forma, serão obtidas as taxas centrais de mortalidade a partir das taxas de mortalidade ( $q_x$ ) da tábua bayesiana, através da seguinte relação:

$$m_x = \frac{2}{(2/q_x)+1} \Rightarrow m_{x,t} = \frac{2}{(2/q_{x,t})+1} \quad (4.8)$$

Uma vez obtidas as taxas centrais de mortalidade  $m_x$ , as mesmas serão projetadas para o ano calendário  $t$  futuro, obtendo-se  $m_{x,t}$  e consequentemente as taxas de mortalidade projetadas  $q_{x,t}$  através da aplicação da relação (4.7). Também, conforme recomendado em Santos (2007), o ideal é que as taxas brutas de mortalidade sejam graduadas antes da obtenção das probabilidades de morte futuras, de forma que as informações sejam suavizadas, produzindo o comportamento de acordo com o esperado (monotônico crescente).

### 4.3 SIMULAÇÃO DO PROCESSO ARIMA(0,1,1) POR *BOOTSTRAP*

Apesar do bom ajuste proporcionado à componente  $k_t$  do modelo de Lee-Carter, através do processo ARIMA(0,1,1); não devemos nos esquecer que os dados utilizados no

ajuste são provenientes da Inglaterra e País de Gales, devido à ausência de um histórico de dados brasileiro relativamente longo, especificamente das populações das EFPC's. Assim, dado que os  $k_t$  foram ajustados com base numa experiência estrangeira, buscou-se avaliar a adequabilidade dos mesmos aos dados da população de expostos ao risco deste estudo.

Para esta avaliação, simulamos valores de  $k_t$  segundo o modelo ARIMA(0,1,1) ajustado, através de *bootstrap* paramétrico, de forma a se obter os  $k_t$  correspondentes aos anos de estudo considerados para a população das EFPC's, isto é,  $k_{2004}, \dots, k_{2008}$ . Para simular um determinado  $k_t$ , primeiramente gera-se o erro aleatório do modelo ARIMA(0,1,1),  $\gamma_t^{(r)}$ , com distribuição normal de média 0 e variância  $\hat{\sigma}_\gamma^2$ . Este procedimento é dado pelos seguintes passos:

1. Gera-se uma observação aleatória  $\hat{\gamma}_t^{(r)} \sim Normal(0, \sigma_\varepsilon^2)$ ;
2. Calcula-se uma diferença  $d_t^{(r)} = \hat{\delta} - \hat{\theta}_1 \gamma_{t-1} + \hat{\gamma}_t^{(r)}$ ;
3. Calcula-se o valor  $k_t^{(r)} = d_t^{(r)} + k_{t-1}^{(r)}$  e retorne ao passo 1.

Neste procedimento,  $r$  indica a  $r$ -ésima iteração, com  $r=1, \dots, 99$  e  $t=2004, \dots, 2008$ , de forma que são gerados 99 valores de  $k_t$  para cada ano de estudo  $t$ . Note que no passo 3,  $k_t^{(r)}$  é calculado somando-se uma diferença  $d_t^{(r)}$  à observação anterior  $k_{t-1}^{(r)}$ . Este artifício é necessário devido à série original ter sido diferenciada uma vez, quando do ajuste do modelo ARIMA(0,1,1).

Obtidos os  $k_t$  simulados para cada ano, utilizaram-se os parâmetros  $a_x$  e  $b_x$  do modelo de Lee-Carter ajustado aos dados da Inglaterra e País de Gales para simular as taxas centrais de mortalidade em cada idade  $x$ , de acordo com a seguinte expressão:

$$m_{x,t}^{(r)} = \exp(\hat{a}_x + \hat{b}_x k_t^{(r)}) \quad (4.9)$$

em que  $x=42, \dots, 90$ ,  $t=2004, \dots, 2008$  e  $r=1, \dots, 99$ .

Assim, para cada ano de estudo  $t$ , obteve-se 49 séries de taxas centrais de mortalidade  $m_{x,t}$ , simuladas de acordo com (4.9), sendo cada série para uma idade  $x$  específica e com tamanho 99. Cada uma destas séries compõe uma linha da matriz  $T$  de taxas centrais de mortalidade simuladas, de dimensão 49 por 99, estruturada da forma:



$$T = \begin{bmatrix} m_{42,t}^1, m_{42,t}^2, \dots, m_{42,t}^{99} \\ m_{43,t}^1, m_{43,t}^2, \dots, m_{43,t}^{99} \\ m_{90,t}^1, m_{90,t}^2, \dots, m_{90,t}^{99} \end{bmatrix}$$

Tomando-se os valores mínimo e máximo em cada linha de  $T$ , obtemos os limites de um intervalo de confiança empírico construído com base em 99 simulações, para as taxas centrais de mortalidade em cada ano  $t$ .

Uma vez construído este intervalo de confiança para cada ano de estudo, é verificado para cada ano de estudo se as taxas centrais de mortalidade da população de expostos ao risco das EFPC's permanecem dentro do intervalo. Uma vez verificada esta condição, conclui-se que os valores dos  $k_t$ , com  $t=2004,2005,\dots,2015$ , projetados pelo modelo ARIMA(0,1,1); são razoáveis na projeção de taxas de mortalidade da população das EFPC's.

A Figura 4.8 apresenta o comportamento das taxas brutas de mortalidade simuladas e o respectivo intervalo de confiança empírico de 99 simulações para o ano de 2008.

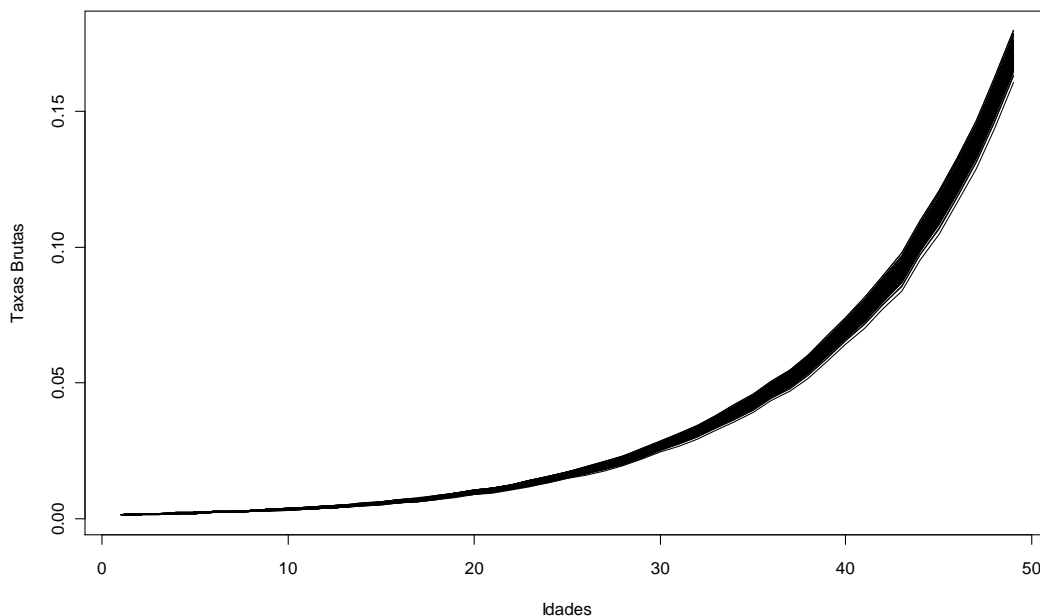


Figura 4.8 – Comportamento de 99 simulações das taxas centrais de mortalidade, em cada idade, considerando o ano de 2008, disponibilizado pelo pacote R.

Verifica-se para os anos de 2005 a 2008, que as taxas centrais de mortalidade das EFPC's se situaram muito próximas aos intervalos de confiança empíricos para algumas

idades e dentro dos intervalos para outras. Embora graficamente tais taxas pareçam estar situadas entre os limites superior e inferior dos intervalos de confiança para a maioria das idades, a Tabela 4.2 aponta as faixas etárias nas quais isto ocorreu de fato, segundo o ano de estudo.

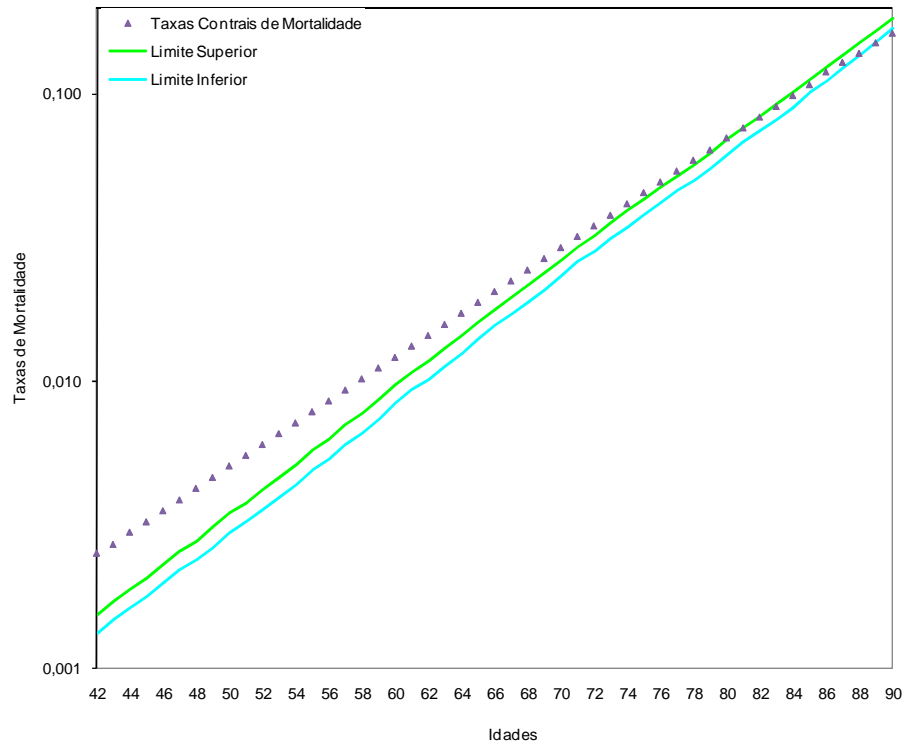


Figura 4.9 – Comparativo das taxas centrais de mortalidade da população das EFPC's em estudo, com aquelas provenientes do intervalo de confiança empírico construído a partir de 99 simulações para o ano de 2005, em escala logarítmica.

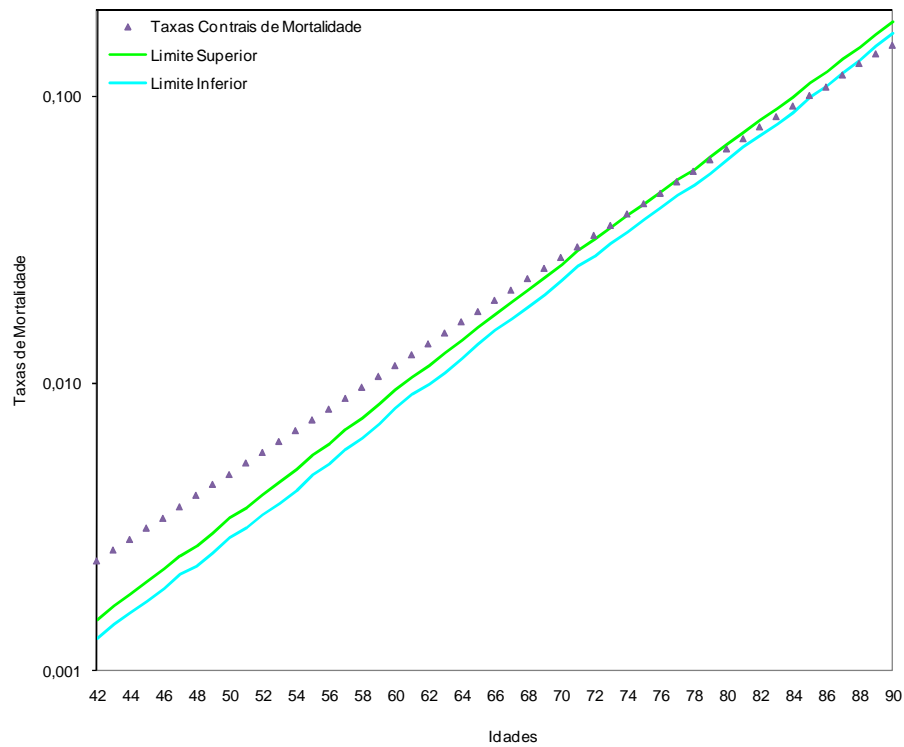


Figura 4.10 – Comparativo das taxas centrais de mortalidade da população das EFPC’s em estudo, com aquelas provenientes do intervalo de confiança empírico construído a partir de 99 simulações para o ano de 2006, em escala logarítmica.

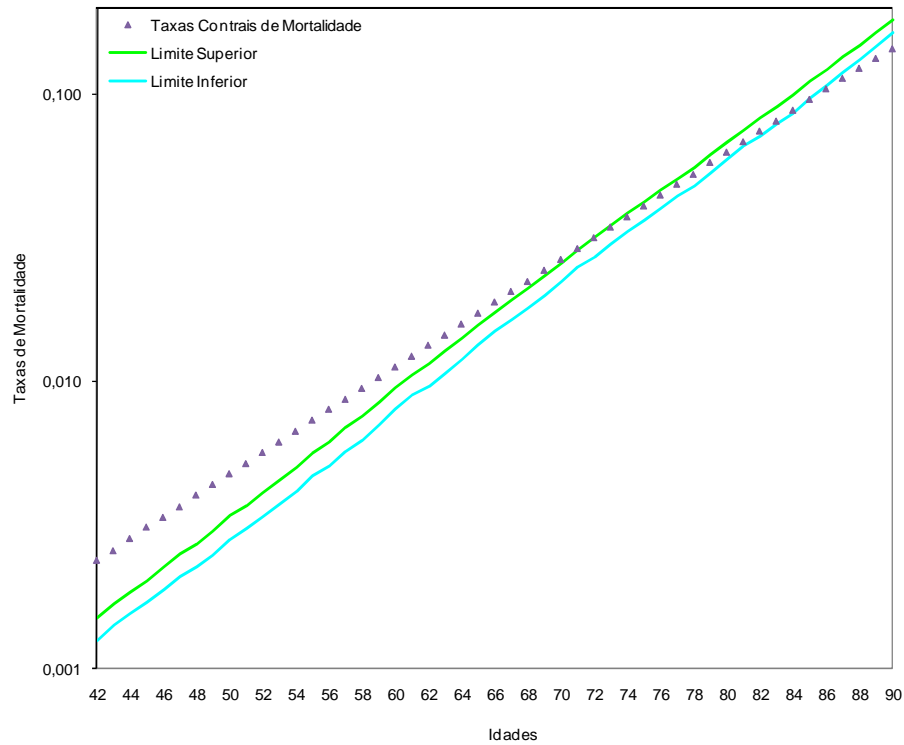


Figura 4.11 – Comparativo das taxas centrais de mortalidade da população das EFPC’s em estudo, com aquelas provenientes do intervalo de confiança empírico construído a partir de 99 simulações para o ano de 2007, em escala logarítmica.

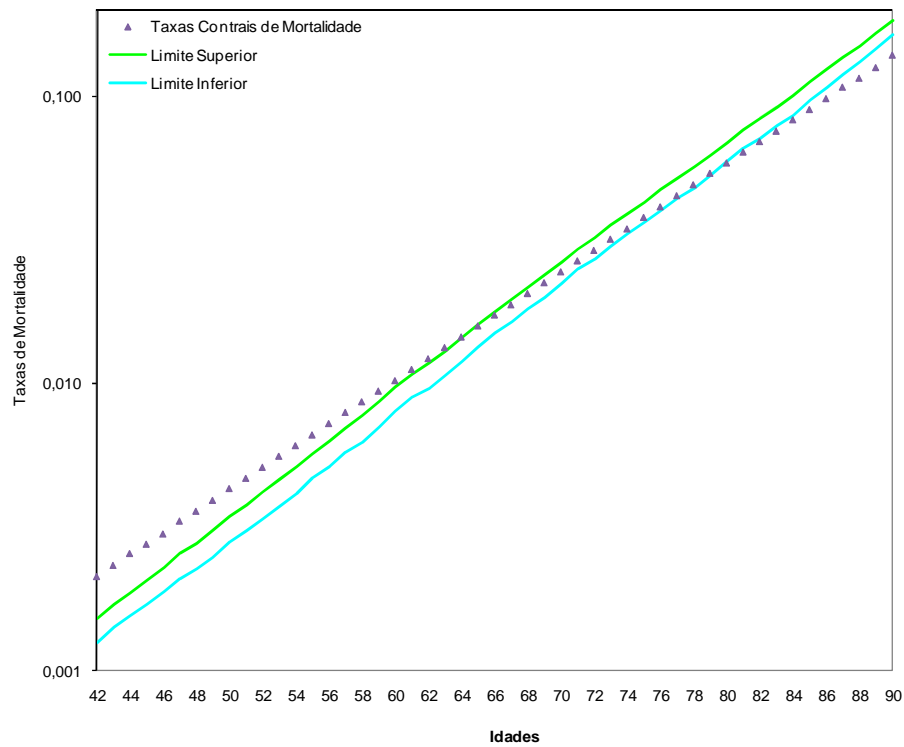


Figura 4.12 – Comparativo das taxas centrais de mortalidade da população das EFPC's em estudo, com aquelas provenientes do intervalo de confiança empírico construído a partir de 99 simulações para o ano de 2008, em escala logarítmica.

Tabela 4.2 – Idades para as quais as taxas centrais de mortalidade se situaram entre os limites superior e inferior dos intervalos de confiança empíricos de 95%, segundo o ano de estudo.

	Ano de Estudo			
	2005	2006	2007	2008
Idades	82 a 89	75 a 85	72 a 84	65 a 79

Como se pode observar da Tabela 4.2, à medida que se avança nos anos, as faixas etárias nas quais as taxas centrais de mortalidade se situam entre os limites de confiança do intervalo empírico com base em 99 simulações, vão se tornando mais amplas. No ano de 2005 apenas 8 idades (82 a 89 anos) foram abrangidas pelos limites de confiança do intervalo, enquanto que em 2008 este total é de 15 idades. Além do aumento da amplitude dessas faixas etárias com o passar dos anos, as idades também vão se tornando cada vez mais jovens, de forma que no ano de 2008 a faixa etária vai de 65 a 79 anos. Logo, neste último ano de estudo, com base no qual a tábua de mortalidade bayesiana foi construída, o modelo de Lee-Carter ajustado à experiência da Inglaterra e País de Gales no período de 1950 a 2003 parece ser razoável, porém para as idades entre aproximadamente 60 e 84 anos, dado que para as idades abaixo de 60 anos as taxas centrais de mortalidade se situam bem acima do limite

superior de confiança, ao passo que acima de 84 anos se situam bem abaixo do limite inferior de confiança, conforme mostra a Figura 4.12.

Portanto, considerando a simulação realizada no último ano de estudo, a componente  $k_t$  do modelo de Lee-Carter ajustado à experiência inglesa, modelada pelo processo ARIMA(0,1,1), gerou taxas centrais de mortalidade levemente inferiores às taxas centrais extraídas da tábua de mortalidade bayesiana para a faixa etária entre 60 e 64 anos e levemente superiores a para faixa etária entre 80 e 84 anos, ao passo que para as idades entre 65 e 79 anos as taxas geradas pelo modelo abrangem aquelas extraídas da tábua bayesiana através do intervalo de confiança empírico. Logo, pode-se dizer que o modelo de Lee-Carter ajustado à experiência inglesa e País de Gales, é coerente na projeção das taxas de mortalidade da população, porém entre 60 e 84 anos de idade.

#### 4.4 PROJEÇÃO DAS TAXAS DE MORTALIDADE DA TÁBUA BAYESIANA

Com base nas considerações feitas na subseção anterior, realizamos as projeções das taxas de mortalidade da tábua bayesiana para os anos de 2009 a 2015, a partir de 2008, utilizando os valores contidos na tabela 4.1. Todavia, restringimos as idades ao intervalo de 60 a 84 anos, devido aos motivos já expostos. A razão da projeção se dar a partir do ano de 2008, é devido à tábua bayesiana já ter contemplado este período em sua construção.

Assim, utilizando (4.6), tem-se que a projeção é dada por:

$$\hat{m}_{x,2008+n} = m_{x,2008} \cdot \exp\left[\hat{b}_x (\hat{k}_{2008+n} - \hat{k}_{2008})\right] \quad (4.10)$$

Note que na expressão acima é a taxa central de mortalidade  $\hat{m}_{x,t}$  que está sendo projetada. Todavia, através da aplicação de (4.7), obtêm-se as taxas de mortalidade projetadas para o ano calendário futuro  $t$  ( $\hat{q}_{x,t}$ ). As taxas de mortalidade da tábua bayesiana, projetadas a partir de 2008, para os anos de 2009 a 2015, está disponível no Anexo VI.

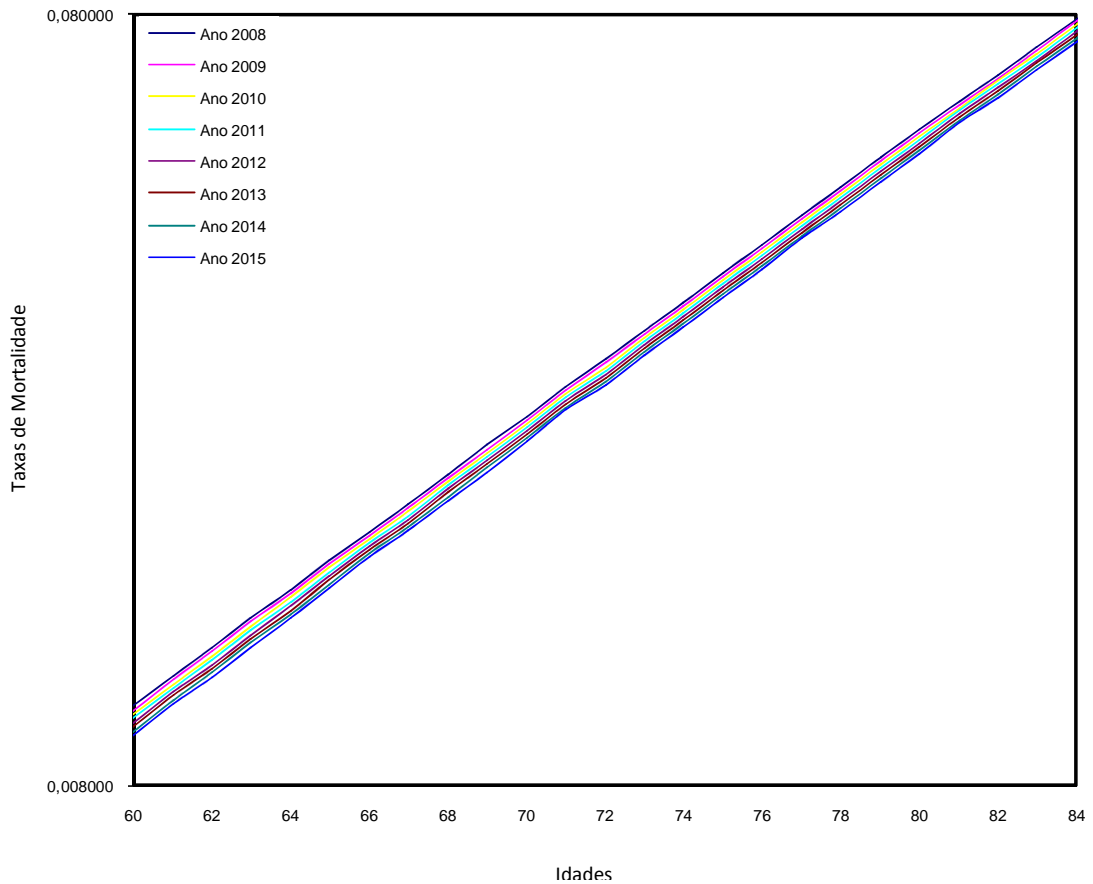


Figura 4.13 – Comportamento das taxas de mortalidade projetadas para os anos de 2009 a 2015, em escala logarítmica, considerando a faixa etária entre 60 e 84 anos.

É nítida a evolução das taxas de mortalidade bayesianas para patamares menores após a aplicação dos fatores de melhoria, conforme (4.10). A Figura 4.13 permite uma melhor visualização do comportamento das taxas de mortalidade na faixa entre 60 e 84 anos, considerando as projeções realizadas. Todavia, os impactos nas Reservas (Provisões) Matemáticas, decorrentes da aplicação das taxas de mortalidade da tábua bayesiana projetadas, serão avaliados a seguir.

#### 4.4.1 IMPACTOS NAS RESERVAS MATEMÁTICAS

O objetivo principal da construção da tábua de mortalidade bayesiana foi obter probabilidades de morte que reflitam a real mortalidade da população de expostos ao risco neste estudo, podendo ser utilizadas em cálculos atuariais, principalmente no dimensionamento das Reservas Matemáticas. Estas reservas representam os compromissos de

uma EFPC com os seus participantes, de forma que se forem subavaliadas, poderá haver desequilíbrio financeiro-atuarial no plano de benefícios previdenciários da mesma. Logo, a utilização de uma tábua de mortalidade que seja uma boa representação da mortalidade da população exposta ao risco no cálculo das Reservas Matemáticas, contribui para o equilíbrio e solvência da entidade.

Neste estudo, os expostos ao risco são aposentados de EFPC's, as quais pagam rendas vitalícias mensais a estes assistidos. A Reserva Matemática Efetiva necessária a dar cobertura aos pagamentos de uma renda vitalícia 13 vezes durante o ano (inclui 13º), a um assistido de idade  $x$ , desconsiderando a reversão de tal renda em pensão por morte a seus beneficiários, é calculada através da seguinte expressão:

$$R_t = 13 \times B_x \times a_x^{(12)}$$

em que  $R_t$  é a Reserva Matemática Efetiva,  $B_x$  representa a renda mensal paga a um assistido de idade  $x$  e  $a_x^{(12)}$  é uma anuidade vitalícia, todas já definidas na seção 2.1 e 3.3. Para avaliar os impactos decorrentes da aplicação da tábua de mortalidade bayesiana no cálculo das Reservas Matemáticas, será envelhecida a massa de expostos ao risco desde o ano de 2008 até 2015, de forma que as Reservas Matemáticas calculadas no ano de 2008 ( $t=0$ ) são as efetivas, enquanto que aquelas calculadas para os anos posteriores são Reservas Matemáticas Esperadas, uma vez que considera em seu cálculo a “probabilidade” do assistido estar vivo nos próximos anos, que nada mais é do que envelhecê-lo. Portanto, a Reserva Matemática Esperada de um assistido com idade  $x$ , é dada por:

$$\hat{R}_t = (13 \times B_x) \times \frac{l_{x+t}}{l_x} \times a_{x+t}^{(12)} = (13 \times B_x) \times {}_t p_x \times a_{x+t}^{(12)}$$

em que  $\hat{R}_t$  é a Reserva Matemática Esperada,  $l_x$  já foi definido anteriormente na seção 2.1 e  ${}_t p_x$  é a probabilidade de um assistido de idade  $x$  alcançar a idade  $x+t$ , com  $t=1,2,\dots,7$ . Assim, para cada ano entre 2008 e 2015, serão calculados os totais das Reservas Matemáticas,  $\sum R_t$  e  $\sum \hat{R}_t$ , tanto pela tábua de mortalidade AT-1983 como pela bayesiana, comparando o acréscimo ou decréscimo em percentuais decorrentes da substituição da tábua AT-1983 pela bayesiana. A comparação da tábua de mortalidade bayesiana com a tábua AT-1983 se deve ao

fato desta última ter se mostrado aderente à população em estudo, conforme o teste realizado na seção 3.3.

Nesta avaliação, as idades serão restritas à faixa de 60 a 84 anos devido às razões já expostas na seção 4.3, o benefício de cada assistido será considerado igual a R\$ 1,00, uma vez que se trata de informação confidencial e não disponível na base de dados e a anuidade  $a_x^{(12)}$  será calculada considerando a taxa de desconto atuarial de 6% ao ano. A evolução das Reservas Matemáticas será avaliada segundo as duas tábuas de mortalidade mencionadas, sendo que para a tábua bayesiana as taxas de mortalidade são desagravadas a cada ano, mediante a aplicação dos fatores de melhoria conforme (4.10), cuja evolução de tais taxas é mostrada na Figura 4.13.

Após o dimensionamento das respectivas reservas, o crescimento ou decréscimo das mesmas será avaliado pela variação em percentual, dada por:

$$p_{EF} = \left[ \left( \frac{\sum R_t^{AT}}{\sum R_t^b} \right) - 1 \right] \times 100$$

e

$$p_{ES} = \left[ \left( \frac{\sum \hat{R}_t^{AT}}{\sum \hat{R}_t^b} \right) - 1 \right] \times 100$$

em que  $R_t^{AT}$  e  $R_t^b$  são as Reservas Matemáticas calculadas pelas tábuas AT-1983 e bayesiana, respectivamente, com  $p_{EF}$  e  $p_{ES}$  sendo as respectivas variações em percentuais das Reservas Matemáticas Efetivas e Esperadas, de forma que  $p_{EF}$  é calculado exclusivamente para o ano de 2008, no qual as Reservas Matemáticas são as efetivas, ao passo que nos anos subsequentes será utilizado somente  $p_{ES}$ , por se tratar de Reservas Matemáticas Esperadas, isto é, da massa de expostos ao risco já envelhecida.

A Tabela 4.3 apresenta a evolução das Reservas Matemáticas.



Tabela 4.3 – Comparativo das evoluções das Reservas Matemáticas segundo as tábuas de mortalidade AT-1983 e bayesiana, para as idades de 60 a 84 anos.

Evolução das Reservas Matemáticas			
<i>t</i>	Tábua AT-1983	Tábua Bayesiana	Variação
2008	2.362.290,46	2.328.199,06	1,46%
2009	2.260.025,67	2.229.781,65	1,36%
2010	2.156.042,69	2.130.218,92	1,21%
2011	2.050.554,90	2.029.663,19	1,03%
2012	1.943.803,46	1.928.285,12	0,80%
2013	1.836.056,05	1.826.267,93	0,54%
2014	1.727.605,04	1.723.804,27	0,22%
2015	1.618.885,01	1.621.227,43	-0,14%

Como se pode observar da Tabela 4.3, entre os anos de 2008 e 2014, a tábua AT-1983 apresentou totais de Reservas Matemáticas superiores aos totais calculados pela tábua bayesiana, uma vez que esta última é menos conservadora que a AT-1983 entre 42 e 79 anos de idade, por apresentar taxas de mortalidade um pouco superiores. Todavia, a partir de 80 anos de idade, a tábua bayesiana se torna mais conservadora que a AT-1983, pois apresenta taxas de mortalidade inferiores. Este comportamento pode ser visualizado na Figura 3.25.

Também é perceptível da Tabela 4.3 que os percentuais de variação vão diminuindo à medida que os anos passam. Este fato é devido aos desagravamentos anuais das taxas de mortalidade da tábua bayesiana, mediante aplicação dos fatores de melhoria conforme (4.10). Isto faz com que essa tábua se torne mais conservadora à medida que os anos passam, ao passo que as taxas de mortalidade da tábua AT-1983 permanecem constantes, isto é, a tábua é estática. Assim, no ano de 2008, as Reservas Matemáticas calculadas através da tábua AT-1983 é 1,46% maior que aquelas calculadas pela tábua bayesiana, diminuindo para 0,22% em 2014. Todavia, em 2015 a situação é inversa, isto é, a tábua bayesiana passa a ser mais conservadora que a AT-1983, de forma que o total das reservas calculadas pela tábua bayesiana é superior ao total calculado pela AT-1983 em 0,14%.

Portanto, comparando-se as duas tábuas de mortalidade para a faixa etária entre 60 e 84 anos, a adoção da tábua AT-1983 como hipótese de mortalidade geral para a população em estudo, é mais prudente até o ano de 2014, sendo que a partir de 2015 a situação é inversa, isto é, a adoção da tábua bayesiana passa a ser mais prudente. Apesar da tábua de mortalidade bayesiana ser mais conservadora para as idades acima de 79 anos, o modelo de Lee-Carter ajustado à experiência inglesa e País de Gales não pareceu razoável na projeção das taxas de

mortalidade acima de 84 anos, pelos motivos já expostos na seção 4.3. Desta forma, os impactos nas Reservas Matemáticas não foram avaliados para idades acima de 84 anos.

## 5 CONCLUSÃO

Este estudo permitiu a investigação dos níveis de mortalidade de uma faixa etária específica - 42 a 90 anos – uma vez que a base de dados não apresentou quantidade suficiente de expostos ao risco fora desta faixa, ainda que tenha sido composta de 6 EFPC's. Mesmo dentro deste intervalo de idades, houve algumas restrições. Através do Modelo Estático Global, escolhido para a construção da tábua de mortalidade bayesiana, graduaram-se todas as taxas de mortalidade dentro da referida faixa etária. Todavia, para as idades a partir de 85 anos, aumenta-se bastante os erros padrões das estimativas das probabilidades de morte ( $\hat{q}_x$ ), diminuindo-se a precisão destas estimativas até 90 anos de idade, conforme mostra o intervalo de confiança de 95% na Figura 3.26. Por esta razão, aceitou-se como coerentes somente as taxas graduadas até 84 anos de idade, devido à grande imprecisão a partir de 85 anos causada pela quantidade bastante reduzida de expostos ao risco nestas idades. Como uma abordagem alternativa, visando contornar este problema, também foram tentados ajustes através do modelo binomial ao invés do modelo de Poisson, porém os resultados obtidos foram muito semelhantes àqueles alcançados pelo modelo Poisson, com exceção de que o modelo binomial não proporcionou a mesma suavidade no crescimento das taxas e nem monotonicidade em algumas idades mais avançadas, razão pela qual preferiu-se o modelo Poisson.

Portanto, considerando as mencionadas restrições, o objetivo da construção da tábua de mortalidade bayesiana foi atingido, podendo avaliar os níveis de mortalidade da população em estudo e obtendo-se um comparativo destes níveis com aquelas tábuas praticadas nas EFPC's, principalmente a AT-1983, cujos impactos nas Reservas Matemáticas foram avaliados.

No que diz respeito à projeção das taxas de mortalidade, o modelo de Lee-Carter ajustado aos dados da Inglaterra e País de Gales é coerente na projeção das taxas de mortalidade da população exposta ao risco, porém para a faixa etária entre 60 e 84 anos. Abaixo de 60 anos e acima de 84 anos, as taxas centrais de mortalidade geradas pelo modelo se distanciam daquelas graduadas para a população em estudo, conforme mostra o intervalo de confiança empírico na Figura 4.12, de forma que as projeções considerando essas idades implicariam em taxas de mortalidade não condizentes com a realidade da população exposta ao risco.

Da realização do teste de aderência na seção 3.3, bem como a avaliação dos impactos nas Reservas Matemáticas, concluímos que a adoção da tábua AT-1983 como hipótese de mortalidade geral, é mais coerente que a tábua bayesiana, porém restrita à faixa de 42 a 79 anos de idade, em que a tábua bayesiana é menos conservadora. Estes resultados também reforçam a determinação da Resolução CGPC N° 18, a qual estabelece a tábua AT-1983 como parâmetro mínimo de mortalidade geral nas EFPC's. Todavia, para as idades acima de 79 anos, em que a tábua bayesiana é mais conservadora, somente foram avaliados os impactos nas reservas matemáticas até 84 anos, devido aos motivos já expostos. Finalmente, os resultados obtidos neste estudo ficaram restritos a faixa entre 42 e 84 anos para efeitos das taxas de mortalidade graduadas, enquanto que para a projeção de tais taxas, a restrição é de 60 a 84 anos de idade.

Para que seja possível graduar taxas de mortalidade fora da amplitude etária considerada, é necessária uma maior quantidade de expostos ao risco, proporcionando maior contingente nas idades em que a exposição foi muito baixa. Dependendo desta quantidade, também pode ser possível a construção de tábuas de mortalidade para cada sexo. Quanto à projeção das taxas de mortalidade, o ideal seria ajustar o modelo de Lee-Carter com base em dados de populações específicas de EFPC's, contudo inexistente um histórico suficientemente longo (em anos de observação) destas populações, impossibilitando um ajuste baseado na realidade dos fundos de pensão brasileiros. Portanto, o fato de não possuímos históricos de dados que remontam a períodos passados relativamente longos e nem quantidade de dados suficientes para alguns seguimentos, ainda consiste num limitador que as vezes impede de nos basear somente em nossa experiência para o desenvolvimento de estudos sobre mortalidade.

Como recomendação para trabalhos futuros, sugere-se a ampliação e aplicação deste estudo, porém com uma massa de expostos ao risco de tamanho maior. É também desejável uma amplitude etária maior, possibilitando também a graduação das taxas de mortalidade segregadas por sexo.

**REFERÊNCIAS**

- BELTRÃO, K.I.; CAMARANO, A.A. **A Dinâmica Populacional Brasileira e a Previdência Social: Uma Descrição com Ênfase nos Idosos**. Relatórios Técnicos 01/99. ENCE/IBGE, Rio de Janeiro, 1999.
- BICKEL, P.J.; DOKSUM, K.A. **Mathematical Statistics**. Prentice Hall, New Jersey, 2007.
- BOOTH, H.; HYNDMAN, R.J.; TICKLE, L.; DE JONG, P. **Lee-Carter mortality forecasting: a multi-country comparison of variants and extensions**. Demography Research, 15(9), 289-310, 2006.
- BOWERS, N.L. et al. **Actuarial Mathematics**. The Society of Actuaries, 1997.
- BOX, G.E.P.; JENKINS, G. M. **Time Series Analysis: Forecasting and Control**. San Francisco: Holden-Day, 1970.
- BROFFIT, J.D. **Increasing and Increasing Convex Bayesian Graduation**. Transactions of Society of Actuaries, v. 40, pp. 115-148, 1988.
- CMB. **Projecting future mortality: A discussion paper**. Continuous Mortality Investigation sub-committee, Working Paper 3, 2004.
- COELHO, E.I.F. **O Método Lee-Carter para Previsão da Mortalidade**. INE, Revista de Estudos Demográficos, 37, 2005.
- COLES, S. **Statistical Inference (Bayesian Inference)**. Lecture Notes, University of Lancaster, 1997.
- COLES, S.; ROBERTS, G.; JARNER, S. **Computer Intensive Methods**. Lecture Notes, University of Lancaster, 2002.
- DeGROOT, M.H. **Probability and Statistics**. Addison-Wesley, Menlo Park, 1989.
- DIGGLE, P.J. **Statistical Analysis of Spatial Point Patterns**. Arnold, London, 2003.
- DOBSON, A.J. **An Introduction to Generalized Linear Models**. Chapman & Hall, London, 1994.
- EHLERS, R.S. **Métodos Computacionalmente Intensivos em Estatística**. Notas de Aula, Universidade Federal do Paraná, 2004.
- EHLERS, R.S. **Introdução à Inferência Bayesiana**. Notas de Aula, Universidade Federal do Paraná, 2006.
- GEMAN, S.; GEMAN, D. **Stochastic Relaxation, Gibbs distributions and the Bayesian restoration of images**. IEEE Trans. Pattn. Anal. Mach. Intel., 6, 721-741, 1984.

GIROSKI, F.; KING, G. **Understanding the Lee-Carter Mortality Forecasting Method**. Harvard University, 2007. Disponível em <http://gking.harvard.edu/files/smooth.pdf>.

HABERMAN, S.; RENSHAW, A. E. **Generalized Linear Models and Actuarial Science**. The Statistician, v. 45, n. 4, pp. 407-436, 1996.

HASTINGS, W.K. **Monte Carlo sampling methods using Markov chains and their applications**. Biometrika, 57, 97-109, 1970.

LEE, R.D.; CARTER, L. **Modeling and Forecasting the Time Series of US Mortality**. Journal of the American Statistical Association, v87, 419: 659-671, 1992.

MARRIOTT, F.H.C. **Barnard's Monte Carlo tests: How many simulations?** Applied Statistics 28, 75-77, 1979.

MONTELLO, J.M. **Notas Técnicas Atuariais**. Jesse Montello Serviços Técnicos em Atuária e Economia – Rio de Janeiro, 2009.

NEVES, C.R. **Graduação Bayesiana de Taxas de Mortalidade**. Funenseg, Caderno de Seguros – Teses, v.10, 28, 2005.

MIGON, H.S.; GAMERMAN, D. **Statistical Inference: an Integrated Approach**. Arnold, London, 1999.

PLUMMER, M.; BEST, N.; COWLES, K.; VINES, K. **coda: Output analysis and diagnostics for MCMC**. R package version 0.13-4. 2009.

PLUMMER, M. **JAGS Version 1.0.3 Manual**. 2009, Disponível em [http://calvin.iarc.fr/~martyn/software/jags/jags\\_user\\_manual.pdf](http://calvin.iarc.fr/~martyn/software/jags/jags_user_manual.pdf).

R Development Core Team (2009). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, 2009.

SANTOS, R.R. **Técnicas de modelagem do improvement para construção de tábuas geracionais**. Pontifícia Universidade Católica do Rio de Janeiro, Dissertação de Mestrado, 2007.

SPIEGELHALTER, D.J.; RICHARDSON, S.; GILKS, W.R. **Markov Chain Monte Carlo in Practice**. Chapman & Hall, London, 1997.

SPIEGELHALTER, D.; THOMAS, A.; BEST, N.; LUNN, D. **WinBUGS User Manual Version 1.4**. Cambridge University, 2003. Disponível em <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/manual14.pdf>

STURTZ, S.; LIGGES, U.; GELMAN, A. **R2WinBUGS: A Package for Running WinBUGS from R**. Journal of Statistical Software, 12(3), 1-16, 2005.

VENABLES, W.N.; RIPLEY, B.D. **Modern Applied Statistics with S**. Springer, New York, 2002.

WATSON WYATT WORLDWIDE, CSIS. **Envelhecimento Mundial – Desafio do Novo Milênio**. Periódico interno, 1999.

WONG, L.L.R.; CARVALHO, J.A. CSIS. **O rápido processo de envelhecimento populacional do Brasil: sérios desafios para as políticas públicas**. São Paulo: Revista Brasileira de Estudos Populacionais, v23, 5-26, 2006.

## ANEXO I – Quantidade Central de Expostos ao Risco

Idade	Ano 2004			Ano 2005			Ano 2006			Ano 2007			Ano 2008		
	M	F	U	M	F	U	M	F	U	M	F	U	M	F	U
42	232	75	307	157	60	217	113	23	136	52	15	67	21	12	33
43	342	157	499	232	75	307	157	60	217	113	23	136	52	15	67
44	483	191	674	337	157	494	231	75	306	157	60	217	112	23	135
45	768	210	978	485	193	678	340	157	497	232	76	308	159	60	219
46	959	262	1221	765	210	975	485	193	678	338	157	495	231	77	308
47	1128	275	1403	958	264	1222	766	211	977	485	195	680	338	157	495
48	1321	262	1583	1129	277	1406	954	268	1222	769	220	989	481	195	676
49	1452	263	1715	1319	268	1587	1130	283	1413	969	283	1252	773	227	1000
50	1438	256	1694	1439	267	1706	1334	274	1608	1139	297	1436	979	295	1274
51	1552	250	1802	1438	258	1696	1446	273	1719	1343	285	1628	1153	308	1461
52	1524	203	1727	1554	251	1805	1441	260	1701	1466	284	1750	1356	296	1652
53	1522	205	1727	1526	205	1731	1561	252	1813	1464	265	1729	1486	290	1776
54	1451	168	1619	1525	208	1733	1548	206	1754	1609	259	1868	1495	269	1764
55	1718	183	1901	1714	216	1930	1812	253	2065	1883	262	2145	1914	327	2241
56	1681	183	1864	1765	190	1955	1737	217	1954	1873	266	2139	1941	272	2213
57	1486	156	1642	1672	182	1854	1768	190	1958	1754	223	1977	1880	270	2150
58	1554	127	1681	1486	158	1644	1671	182	1853	1783	196	1979	1763	223	1986
59	1437	115	1552	1550	128	1678	1474	158	1632	1674	186	1860	1785	195	1980
60	1374	93	1467	1425	122	1547	1532	129	1661	1469	164	1633	1658	188	1846
61	1222	88	1310	1360	98	1458	1425	124	1549	1518	134	1652	1460	165	1625
62	1029	90	1119	1201	89	1290	1346	98	1444	1415	127	1542	1506	134	1640
63	932	82	1014	1010	90	1100	1184	88	1272	1343	99	1442	1410	127	1537
64	860	56	916	908	82	990	994	91	1085	1167	89	1256	1320	99	1419
65	753	63	816	844	57	901	893	81	974	990	93	1083	1148	88	1236
66	646	42	688	737	64	801	830	55	885	878	82	960	982	92	1074
67	591	38	629	639	42	681	725	65	790	817	54	871	861	81	942
68	500	49	549	565	37	602	623	42	665	717	64	781	805	54	859
69	457	34	491	486	49	535	543	37	580	611	42	653	699	64	763
70	332	22	354	447	34	481	471	49	520	537	37	574	598	41	639
71	312	29	341	321	22	343	431	33	464	459	48	507	527	36	563
72	240	18	258	302	29	331	302	22	324	414	33	447	447	48	495
73	226	14	240	231	18	249	288	29	317	293	21	314	405	33	438
74	195	19	214	215	14	229	219	18	237	276	29	305	283	21	304
75	165	13	178	187	19	206	209	13	222	213	18	231	262	29	291
76	121	12	133	161	13	174	178	19	197	202	12	214	205	18	223
77	112	10	122	118	12	130	155	13	168	169	19	188	190	12	202
78	102	8	110	105	10	115	110	12	122	150	13	163	158	19	177
79	55	6	61	95	8	103	100	9	109	103	12	115	143	13	156
80	49	3	52	48	6	54	92	7	99	96	9	105	97	12	109
81	56	8	64	47	3	50	46	5	51	87	7	94	88	8	96
82	51	6	57	53	7	60	45	3	48	41	5	46	78	7	85
83	23	2	25	45	6	51	48	7	55	42	3	45	37	5	42
84	18	2	20	20	2	22	41	4	45	45	6	51	41	3	44
85	13	3	16	16	2	18	19	2	21	37	4	41	42	6	48
86	9	0	9	13	3	16	15	1	16	18	2	20	33	4	37
87	12	0	12	9	0	9	9	3	12	14	0	14	15	2	17
88	7	1	8	9	0	9	8	0	8	8	3	11	12	0	12
89	2	0	2	7	1	8	5	0	5	8	0	8	7	3	10
90	1	0	1	1	0	1	6	1	7	5	0	5	7	0	7

M: masculino; F: feminino; U: unissex



## ANEXO II – Quantidade de Mortes Observadas

Idade	Ano 2004			Ano 2005			Ano 2006			Ano 2007			Ano 2008		
	M	F	U	M	F	U	M	F	U	M	F	U	M	F	U
42	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
43	5	0	5	1	0	1	0	0	0	1	0	1	0	0	0
44	3	1	4	1	1	2	1	0	1	0	0	0	0	0	0
45	5	0	5	1	0	1	2	0	2	2	0	2	0	0	0
46	3	0	3	2	0	2	2	0	2	1	0	1	2	0	2
47	4	1	5	7	0	7	4	0	4	5	1	6	1	0	1
48	6	0	6	4	1	5	0	0	0	3	0	3	1	0	1
49	16	1	17	2	2	4	6	0	6	6	0	6	5	1	6
50	8	0	8	4	0	4	8	1	9	4	1	5	4	0	4
51	9	1	10	6	0	6	6	0	6	8	0	8	4	0	4
52	12	1	13	10	0	10	4	0	4	6	0	6	5	0	5
53	9	0	9	11	0	11	7	0	7	9	0	9	8	0	8
54	13	0	13	7	0	7	4	0	4	7	0	7	5	0	5
55	14	1	15	13	1	14	12	0	12	12	0	12	15	0	15
56	13	1	14	8	1	9	19	0	19	12	0	12	9	0	9
57	9	0	9	12	0	12	18	0	18	18	1	19	17	0	17
58	10	0	10	15	0	15	12	0	12	10	2	12	19	1	20
59	16	0	16	20	0	20	11	0	11	21	0	21	15	0	15
60	16	0	16	7	0	7	21	0	21	17	1	18	15	0	15
61	23	0	23	14	0	14	14	0	14	18	0	18	16	0	16
62	21	0	21	17	1	18	6	1	7	10	0	10	12	0	12
63	27	1	28	16	0	16	18	0	18	26	0	26	23	0	23
64	18	0	18	16	1	17	5	0	5	23	1	24	21	0	21
65	20	0	20	15	3	18	19	0	19	12	1	13	27	0	27
66	8	0	8	12	1	13	17	1	18	18	1	19	18	1	19
67	26	1	27	16	0	16	12	1	13	12	0	12	12	2	14
68	15	0	15	22	0	22	13	0	13	18	0	18	23	1	24
69	10	0	10	15	0	15	7	0	7	13	1	14	21	2	23
70	12	0	12	16	1	17	13	1	14	11	1	12	16	0	16
71	10	0	10	19	0	19	17	0	17	13	0	13	17	1	18
72	9	0	9	14	0	14	9	1	10	9	0	9	12	2	14
73	11	0	11	13	0	13	12	0	12	11	0	11	13	1	14
74	8	0	8	6	1	7	6	0	6	14	0	14	13	0	13
75	5	0	5	9	0	9	7	1	8	8	0	8	12	0	12
76	3	0	3	6	0	6	9	0	9	12	0	12	11	0	11
77	7	0	7	8	0	8	5	0	5	11	0	11	6	0	6
78	7	0	7	5	1	6	7	0	7	7	0	7	12	0	12
79	7	0	7	3	1	4	4	0	4	6	0	6	7	1	8
80	2	0	2	2	1	3	5	0	5	8	1	9	3	1	4
81	3	1	4	2	0	2	5	0	5	9	0	9	3	0	3
82	6	0	6	5	0	5	3	0	3	4	0	4	4	0	4
83	2	0	2	4	2	6	3	1	4	1	0	1	3	1	4
84	2	0	2	1	0	1	4	0	4	3	0	3	8	0	8
85	0	0	0	1	1	2	1	0	1	4	0	4	3	0	3
86	0	0	0	4	0	4	1	1	2	3	0	3	4	1	5
87	3	0	3	1	0	1	1	0	1	2	0	2	0	0	0
88	0	0	0	4	0	4	0	0	0	1	0	1	2	0	2
89	1	0	1	1	0	1	0	0	0	1	0	1	1	1	2
90	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

M: masculino; F: feminino; U: unissex





## ANEXO IV – Tábua de Mortalidade Bayesiana Construída sob o Modelo Estático Local

### A) Tábua de Mortalidade Bayesiana Estática

Idade	Prob. de Morte Estimada	Intervalo de Probabilidade	
		2,50%	97,50%
42	0,002123	0,000000	0,006472
43	0,002324	0,000000	0,006018
44	0,002554	0,000656	0,005239
45	0,002752	0,000885	0,005296
46	0,002983	0,001258	0,005022
47	0,003309	0,001653	0,005185
48	0,003580	0,002059	0,005233
49	0,003905	0,002477	0,005565
50	0,004291	0,002844	0,005814
51	0,004654	0,003224	0,006191
52	0,005085	0,003634	0,006672
53	0,005550	0,003993	0,007176
54	0,006038	0,004407	0,007784
55	0,006580	0,004998	0,008252
56	0,007207	0,005494	0,008853
57	0,007878	0,006157	0,009672
58	0,008579	0,006695	0,010610
59	0,009336	0,007374	0,011370
60	0,010140	0,008056	0,012410
61	0,011090	0,008869	0,013520
62	0,012080	0,009563	0,014850
63	0,013200	0,010580	0,016050
64	0,014360	0,011330	0,017440
65	0,015680	0,012360	0,019030
66	0,017080	0,013400	0,020970
67	0,018530	0,014430	0,022990
68	0,020280	0,015780	0,025180
69	0,022080	0,017070	0,027470
70	0,024060	0,018290	0,029940
71	0,026250	0,019720	0,032920
72	0,028520	0,021270	0,035840
73	0,031140	0,023530	0,039670
74	0,033760	0,024780	0,043470
75	0,036960	0,026640	0,048570
76	0,040180	0,028140	0,052690
77	0,043850	0,029950	0,057900
78	0,047720	0,032260	0,063470
79	0,052190	0,034560	0,071050
80	0,056550	0,037730	0,077990
81	0,061590	0,039010	0,086240
82	0,066730	0,040430	0,096600
83	0,072370	0,039680	0,107200
84	0,078940	0,045790	0,117500
85	0,085450	0,044160	0,132300
86	0,092960	0,041830	0,149900
87	0,101400	0,038700	0,168200
88	0,108400	0,035710	0,196000
89	0,117600	0,025970	0,231400
90	0,129000	0,000000	0,273900

### B) Tábua de Mortalidade Bayesiana Geracional

Idade	$\hat{q}_{x, 2008}$	$\hat{q}_{x, 2009}$	$\hat{q}_{x, 2010}$	$\hat{q}_{x, 2011}$	$\hat{q}_{x, 2012}$	$\hat{q}_{x, 2013}$	$\hat{q}_{x, 2014}$	$\hat{q}_{x, 2015}$
42	0,002123	0,002096	0,002070	0,002044	0,002019	0,001993	0,001968	0,001944
43	0,002324	0,002296	0,002269	0,002242	0,002216	0,002190	0,002164	0,002138
44	0,002554	0,002523	0,002493	0,002462	0,002433	0,002403	0,002374	0,002345
45	0,002752	0,002717	0,002683	0,002649	0,002616	0,002583	0,002550	0,002518
46	0,002983	0,002945	0,002908	0,002871	0,002835	0,002799	0,002764	0,002729
47	0,003309	0,003267	0,003225	0,003183	0,003143	0,003102	0,003062	0,003023
48	0,003580	0,003532	0,003485	0,003439	0,003393	0,003348	0,003304	0,003260
49	0,003905	0,003851	0,003798	0,003746	0,003694	0,003643	0,003593	0,003544
50	0,004291	0,004234	0,004178	0,004123	0,004068	0,004014	0,003961	0,003908
51	0,004654	0,004593	0,004533	0,004474	0,004415	0,004357	0,004300	0,004244
52	0,005085	0,005015	0,004946	0,004878	0,004811	0,004744	0,004679	0,004615
53	0,005550	0,005473	0,005398	0,005323	0,005250	0,005177	0,005106	0,005035
54	0,006038	0,005955	0,005872	0,005791	0,005711	0,005632	0,005555	0,005478
55	0,006580	0,006495	0,006412	0,006329	0,006248	0,006168	0,006088	0,006010
56	0,007207	0,007110	0,007015	0,006921	0,006828	0,006736	0,006646	0,006557
57	0,007878	0,007774	0,007672	0,007571	0,007471	0,007373	0,007276	0,007180
58	0,008579	0,008463	0,008349	0,008236	0,008125	0,008015	0,007907	0,007800
59	0,009336	0,009214	0,009094	0,008975	0,008858	0,008742	0,008628	0,008515
60	0,010140	0,010015	0,009892	0,009770	0,009650	0,009531	0,009414	0,009298
61	0,011090	0,010960	0,010831	0,010703	0,010577	0,010453	0,010330	0,010209
62	0,012080	0,011927	0,011776	0,011627	0,011480	0,011335	0,011192	0,011050
63	0,013200	0,013034	0,012870	0,012708	0,012548	0,012390	0,012234	0,012080
64	0,014360	0,014183	0,014008	0,013835	0,013665	0,013496	0,013330	0,013165
65	0,015680	0,015495	0,015311	0,015130	0,014951	0,014774	0,014599	0,014427
66	0,017080	0,016893	0,016708	0,016524	0,016343	0,016164	0,015987	0,015811
67	0,018530	0,018318	0,018109	0,017902	0,017698	0,017495	0,017296	0,017098
68	0,020280	0,020052	0,019827	0,019604	0,019384	0,019166	0,018951	0,018738
69	0,022080	0,021832	0,021586	0,021344	0,021104	0,020866	0,020632	0,020400
70	0,024060	0,023803	0,023549	0,023298	0,023050	0,022804	0,022560	0,022319
71	0,026250	0,025993	0,025738	0,025486	0,025236	0,024989	0,024744	0,024501
72	0,028520	0,028215	0,027914	0,027616	0,027321	0,027029	0,026740	0,026455
73	0,031140	0,030809	0,030481	0,030156	0,029835	0,029518	0,029203	0,028892
74	0,033760	0,033402	0,033048	0,032697	0,032350	0,032007	0,031667	0,031331
75	0,036960	0,036578	0,036199	0,035825	0,035454	0,035087	0,034723	0,034364
76	0,040180	0,039758	0,039340	0,038926	0,038517	0,038112	0,037711	0,037314
77	0,043850	0,043413	0,042981	0,042553	0,042129	0,041709	0,041293	0,040882
78	0,047720	0,047224	0,046733	0,046247	0,045766	0,045290	0,044819	0,044352
79	0,052190	0,051648	0,051112	0,050581	0,050055	0,049535	0,049020	0,048511
80	0,056550	0,056000	0,055455	0,054915	0,054380	0,053850	0,053325	0,052806
81	0,061590	0,061023	0,060461	0,059905	0,059353	0,058806	0,058264	0,057727
82	0,066730	0,066083	0,065442	0,064807	0,064177	0,063554	0,062937	0,062325
83	0,072370	0,071671	0,070978	0,070291	0,069612	0,068938	0,068271	0,067610
84	0,078940	0,078154	0,077376	0,076605	0,075842	0,075086	0,074337	0,073595
85	0,085450	0,084678	0,083912	0,083153	0,082400	0,081654	0,080914	0,080181
86	0,092960	0,092111	0,091268	0,090434	0,089606	0,088786	0,087973	0,087167
87	0,101400	0,100528	0,099662	0,098804	0,097953	0,097109	0,096272	0,095441
88	0,108400	0,107528	0,106663	0,105804	0,104952	0,104106	0,103267	0,102434
89	0,117600	0,116689	0,115785	0,114888	0,113997	0,113113	0,112235	0,111363
90	0,129000	0,128088	0,127183	0,126283	0,125389	0,124501	0,123619	0,122743

## ANEXO V – Estimativas dos Parâmetros do Modelo de Lee-Carter

A) Estimativas dos parâmetros  $\hat{a}_x$  e  $\hat{b}_x$  ajustado aos dados da Inglaterra e País de Gales

Idade	$\hat{a}_x$	$\hat{b}_x$	Idade	$\hat{a}_x$	$\hat{b}_x$	Idade	$\hat{a}_x$	$\hat{b}_x$
42	-6,14879	0,02229	59	-4,40257	0,02330	76	-2,77203	0,01902
43	-6,05596	0,02106	60	-4,30594	0,02196	77	-2,69252	0,01805
44	-5,95802	0,02153	61	-4,21746	0,02099	78	-2,59245	0,01889
45	-5,84422	0,02240	62	-4,10548	0,02260	79	-2,49507	0,01891
46	-5,73666	0,02251	63	-4,00911	0,02249	80	-2,40735	0,01776
47	-5,62575	0,02281	64	-3,90686	0,02205	81	-2,32745	0,01683
48	-5,53026	0,02365	65	-3,81651	0,02117	82	-2,21583	0,01779
49	-5,41343	0,02453	66	-3,74020	0,01962	83	-2,12175	0,01778
50	-5,31031	0,02361	67	-3,62549	0,02046	84	-2,02238	0,01838
51	-5,22824	0,02331	68	-3,53672	0,02013	85	-1,93510	0,01674
52	-5,10861	0,02453	69	-3,43754	0,02017	86	-1,83373	0,01699
53	-5,00590	0,02460	70	-3,34871	0,01915	87	-1,75318	0,01606
54	-4,90673	0,02462	71	-3,26855	0,01760	88	-1,67273	0,01507
55	-4,81991	0,02292	72	-3,14989	0,01922	89	-1,57924	0,01457
56	-4,70903	0,02392	73	-3,05345	0,01918	90	-1,49591	0,01338
57	-4,60593	0,02348	74	-2,95601	0,01914			
58	-4,50146	0,02411	75	-2,87140	0,01870			

B) Estimativas de Segundo Estágio dos Parâmetros  $\hat{k}_t$

Ano	$\hat{k}_t$	Ano	$\hat{k}_t$	Ano	$\hat{k}_t$
1950	11,7054	1968	6,1310	1986	-3,3022
1951	15,0404	1969	5,8597	1987	-5,1162
1952	9,9835	1970	4,8600	1988	-5,5165
1953	9,8966	1971	3,4797	1989	-5,7406
1954	9,1136	1972	5,0120	1990	-7,3453
1955	10,1026	1973	4,1706	1991	-7,5293
1956	9,7422	1974	3,5579	1992	-8,9860
1957	8,2969	1975	2,9581	1993	-7,9653
1958	8,7356	1976	3,7690	1994	-10,6332
1959	8,0601	1977	1,5355	1995	-9,9499
1960	7,1206	1978	1,6699	1996	-11,2559
1961	8,5614	1979	1,6605	1997	-12,2461
1962	8,3875	1980	0,1904	1998	-12,8047
1963	9,0206	1981	-0,6624	1999	-13,3389
1964	4,9684	1982	-0,8486	2000	-15,6499
1965	5,5169	1983	-1,5900	2001	-16,8016
1966	6,1817	1984	-3,2699	2002	-17,2355
1967	3,8125	1985	-1,9817	2003	-17,4928

## APÊNDICE

### A) Modelo de Carlin Modificado Com $\beta$ Fixo (Modelo 1) para WinBUGS

```
#####
# Codigo construido para executar o Modelo 1 no WinBUGS a partir      #
# do R, atraves do pacote R2WinBUGS                                  #
#####

### carregamento do pacote R2WinBUGS e entrada de informacoes
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iteracoes:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento:"))
n.thin=as.numeric(readline("informe o tamanho do intervalo (thin):"))

### Especificacao do diretorio de trabalho e variavel monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","modelo1.txt")
WD="c:\\scripts\\graduacao\\bugs1\\"

### Criacao do arquivo de dados
data=read.table(file=paste(WD,"dados.txt", sep=""),header=T)
n=nrow(data)
d=data$mortes
e=data$expo
data=list("n","d","e")

### Estabelecimento dos valores iniciais das cadeias
ini=read.table(file=paste(WD,"inicio.txt", sep=""), header=T)
inits<-list(list(theta=ini$AT2000),list(theta=ini$AT83),
list(theta=ini$AT49))

### Estabelecimento de parametros para a funcao bugs e execucao do
### WinBUGS, gerando traces das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=3,
n.iter, n.burnin, n.thin,
bugs.directory="c:\\Program Files\\WinBUGS14\\",
program=c("WinBUGS"), debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
  quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
  ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatisticas
### decorrentes da implementacao do modelo e diagnosticos de convergencia
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
plot(out.sim) # para execucao deste linha, codaPkg=F na funcao bugs
```

```
#####
# Codigo para implementacao do Modelo 1 (modelo1.txt) no WinBUGS      #
#  construido no diretorio:                                          #
#  "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"     #
#####

### funcao de verossimilhanca
model{
  for (i in 1:n){
    d[i]~dpois(lambda[i])
    lambda[i]<-theta[i]*e[i]
  }

### Estabelecimento das prioris de theta[i], para i=1,...,n
  theta[1]~dgamma(0.001,0.001)I(0,theta[2])
  for (i in 1:n-1){
    theta[i]~dgamma(0.001,0.001)I(theta[i-1],theta[i+1])
  }
  theta[n]~dgamma(0.001,0.001)I(theta[n-1],1)

### Estabelecimento da distribuicao preditiva do numero de mortes
  for (i in 1:n){
    drep[i]~dpois(lambda[i])
    theta.est[i]<-drep[i]/e[i]
    q.est[i]<-1-exp(-theta.est[i])
  }
}
```



## B) Modelo com Reparametrização tipo Broffit (Modelo 2) para WinBUGS

```
#####
# Codigo construido para executar o Modelo 2 no WinBUGS a partir      #
# do R, atraves do pacote R2WinBUGS                                  #
#####

### carregamento do pacote R2WinBUGS e entrada de informacoes
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iteracoes:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento:"))
n.thin=as.numeric(readline("informe o tamanho do intervalo (thin):"))

### Especificacao do diretorio de trabalho e variavel monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","modelo2.txt")
WD="c:\\scripts\\graduacao\\bugs2\\"

### Carga e criacao do arquivo de dados
data=read.table(file=paste(WD,"dados.txt", sep=""),header=T)
n=nrow(data)
d=data$mortes
e=data$expo
data=list("n","d","e")

### Estabelecimento dos valores iniciais das cadeias
ini=read.table(file=paste(WD,"inicio_phi.txt", sep=""), header=T)
inits<-list(list(phi=ini$AT2000),list(phi=ini$AT83),list(phi=ini$AT49))

### Estabelecimento de parametros para a funcao bugs e execucao do
### WinBUGS, gerando traces das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=3, n.iter,
n.burnin, n.thin, bugs.directory="c:\\Program Files\\WinBUGS14\\",
program=c("WinBUGS"), debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
  quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
  ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatisticas
### decorrentes da implementacao do modelo e diagnosticos de convergencia
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
plot(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
```

```
#####
# Codigo para implementacao do Modelo 2 (modelo2.txt) no WinBUGS      #
#  construido no diretorio:                                          #
#  "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"    #
#####

### funcao de verossimilhanca
model{
  for (i in 1:n){
    lambda[i]<-theta[i]*e[i]
    d[i]~dpois(lambda[i])
  }

### Hiperparametros
  for (i in 1:n){
    beta[i]~dgamma(0.1,0.1)
  }

### Estabelecimento da priori de phi[i], para i=1,...,n
  for (i in 1:n){
    phi[i]~dgamma(2,beta[i])
  }

### Reparametrizacao
  theta[1]<-phi[1]

  for (i in 2:n){
    theta[i]<-theta[i-1]+phi[i]
  }

### Estabelecimento da distribuicao preditiva do numero de mortes
  for (i in 1:n){
    drep[i]~dpois(lambda[i])
    theta.est[i]<-drep[i]/e[i]
    q.est[i]<-1-exp(-theta.est[i])
  }
}
```

### C) Modelo com Força de Mortalidade tipo Gompertz (Modelo 3) para WinBUGS

```
#####
# Codigo construido para executar o Modelo 3 no WinBUGS a partir      #
# do R, atraves do pacote R2WinBUGS                                  #
#####

### carregamento do pacote R2WinBUGS e entrada de informacoes
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iterações:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento:"))
n.thin=as.numeric(readline("Informe o tamanho do intervalo (thin):"))
b=as.numeric(readline("Informe o valor de beta (>0):"))
g=as.numeric(readline("Informe o valor de gama (>=1):"))

### Especificacao do diretorio de trabalho e variavel monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","modelo3.txt")
WD="c:\\scripts\\graduacao\\bugs3\\"

### Carga e criacao do arquivo de dados
data=read.table(file=paste(WD,"dados.txt", sep=""),header=T)
n=nrow(data)
d=data$mortes
e=data$expo # 'e' e um vetor contendo a var. 'expo' da tabela 'data'
data=list("n","d","e")

### Estabelecimento dos valores iniciais das cadeias
inits1<-list(gama=g,beta=b)
inits<-list(inits1)

### Estabelecimento de parametros para a funcao bugs e execucao do
### WinBUGS, gerando traces das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=1, n.iter,
n.burnin, n.thin, bugs.directory="c:\\Program Files\\WinBUGS14\\",
program=c("WinBUGS"), debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatisticas
### decorrentes da implementacao do modelo e diagnosticos de convergencia
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
plot(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
```

```
#####
# Codigo para implementacao do Modelo 3 (modelo3.txt) no WinBUGS      #
#  construido no diretorio:                                          #
#  "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"    #
#####

### funcao de verossimilhanca
model{
  for (i in 1:n){
    log(lambda[i]) <- log(e[i]) + log(beta) + (47+i) * log(gama)
    d[i] ~ dpois(lambda[i])
    theta[i] <- beta * pow(gama, (47+i))
  }

### Estabelecimento das prioris dos parametros beta e gama

    beta ~ dnorm(0, 0.0001) I(0,)
    gama ~ dnorm(0, 0.0001) I(1,)

### Estabelecimento da distribuicao preditiva do numero de mortes
  for (i in 1:n){
    drep[i] ~ dpois(lambda[i])
    theta.est[i] <- drep[i] / e[i]
    q.est[i] <- 1 - exp(-theta.est[i])
  }
}
```

#### D) Modelo com Força de Mortalidade tipo Makeham (Modelo 4) para WinBUGS

```
#####
# Codigo construido para executar o Modelo 4 no WinBUGS a partir      #
# do R, atraves do pacote R2WinBUGS                                  #
#####

### carregamento do pacote R2WinBUGS e entrada de informacoes
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iteracoes:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento:"))
n.thin=as.numeric(readline("Informe o tamanho do intervalo (thin):"))
a=as.numeric(readline("Informe o valor de alpha (>=0):"))
b=as.numeric(readline("Informe o valor de beta (>0):"))
g=as.numeric(readline("Informe o valor de gama (>=1):"))

### Especificacao do diretorio de trabalho e variavel monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","modelo4.txt")
WD="c:\\scripts\\graduacao\\bugs4\\"

### Carga e criacao do arquivo de dados
data=read.table(file=paste(WD,"dados.txt", sep=""),header=T)
n=nrow(data)
d=data$mortes
e=data$expo
data=list("n","d","e")

### Estabelecimento dos valores iniciais das cadeias
inits1<-list(alpha=a,beta=b,gama=g)
inits<-list(inits1)

### Estabelecimento de parametros para a funcao bugs e execucao do
### WinBUGS, gerando traces das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=1, n.iter,
n.burnin, n.thin, bugs.directory="c:\\Program Files\\WinBUGS14\\",
program=c("WinBUGS"), debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatisticas
### decorrentes da implementacao do modelo e diagnosticos de convergencia
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
plot(out.sim) # para execucao desta linha, codaPkg=F na funcao bugs
```

```
#####
# Codigo para implementacao do Modelo 4 (modelo4.txt) no WinBUGS      #
#  construido no diretorio:                                          #
#  "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"    #
#####

### funcao de verossimilhanca
model{
  for (i in 1:n){
    log(lambda[i])<-log(e[i])+log(alpha+beta*pow(gama,(47+i)))
    d[i]~dpois(lambda[i])
    theta[i]<-alpha+beta*pow(gama,(47+i))
  }

### Estabelecimento das prioris dos parametros alpha, beta e gama
alpha~dnorm(0,0.0001)I(0,)
beta~dnorm(0,0.0001)I(0,)
gama~dnorm(0,0.0001)I(1,)

### Estabelecimento da distribuicao preditiva do numero de mortes
  for (i in 1:n){
    drep[i]~dpois(lambda[i])
    theta.est[i]<-drep[i]/e[i]
    q.est[i]<-1-exp(-theta.est[i])
  }
}
```

E) Modelo de Carlin Modificado Com  $\beta$  Fixo (Modelo 1) para JAGS

```
#####
#Codigo construido para executar o Modelo 1 no JAGS a partir do R #
#####

### Carregamento do pacote coda e do scrit jags.r e entrada de informacoes
require(coda) # REQUER O PACOTE CODA
source("c:\\scripts\\graduacao\\modelo1\\jags.r")
mcmc=list()
mcmc$niter=readline("Informe o total de iteracoes:")
mcmc$nburn=readline("Informe o tamanho do aquecimento (burn in):")
mcmc$nthin=readline("informe o tamanho do intervalo (thin):")

### Especificacao da variavel a ser monitorada e criacao do arquivo de
### comandos a ser interpretado pelo JAGS
vars=c('q.est')
jags.cmd(vars=vars,name='model',data='obitos',mcmc=mcmc)

### Carga e criacao do arquivo de dados
data=read.table(file="c:\\scripts\\graduacao\\modelo1\\dados.txt",
header=T)
d=data$mortes
e=data$expo
dump(c('d','e'), file="c:\\scripts\\graduacao\\modelo1\\obitos-data.R")

### Estabelecimento dos valores iniciais das cadeias
ini=read.table(file="c:\\scripts\\graduacao\\modelo1\\inicio.txt",
header=T)
theta=ini$AT2000
dump(c('theta'), file="c:\\scripts\\graduacao\\modelo1\\obitos-par1.R")
theta=ini$AT83
dump(c('theta'), file="c:\\scripts\\graduacao\\modelo1\\obitos-par2.R")
theta=ini$AT49
dump(c('theta'), file="c:\\scripts\\graduacao\\modelo1\\obitos-par3.R")

### Carga do pacote JAGS pelo R para execucao do arquivo comandos.cmd
system('jags comandos.cmd')

### Estabelecimento de parametros para a funcao openbugs e geracao dos
### traces, densidades a posteriori, estimativas e estatisticas
output=read.openbugs(stem="c:\\scripts\\graduacao\\modelo1\\", quiet=F)
plot(output, trace=T, density=T, smooth=T, auto.layout=T,
ask=dev.interactive())
```

```
#####
# Código contido no script jags.r para geracao do arquivo comandos.cmd #
#####

jags.cmd=function(vars,data,name,mcmc,seed) {
  file=paste('comandos.cmd')
  text=NULL
  nvar=length(vars)
  if (!missing(seed)) text=paste(text,"seed ",seed,"\n",sep="")
  text=paste(text,'model in "c:\\scripts\\graduacao\\modelo1\\',name,'.bug
  "\n',sep='')
  text=paste(text,'data in "c:\\scripts\\graduacao\\modelo1\\',data,'-data
  .R"\n',sep='')
  text=paste(text,'compile, nchains(3)\n',sep='')
  text=paste(text,'parameters in "c:\\scripts\\graduacao\\modelo1\\',data,
  '-par1.R", chain(1) \n',sep='')
  text=paste(text,'parameters in "c:\\scripts\\graduacao\\modelo1\\',data,
  '-par2.R", chain(2) \n',sep='')
  text=paste(text,'parameters in "c:\\scripts\\graduacao\\modelo1\\',data,
  '-par3.R", chain(3) \n',sep='')
  text=paste(text,'initialize\n',sep='')
  text=paste(text,'update ',as.integer(mcmc$nburn),'\n',sep='')
  aux=NULL
  if (mcmc$nthin>1) aux=paste(', ',thin('mcmc$nthin,')',sep='')

  for(i in 1:nvar) text=paste(text,'monitor set ',vars[i],aux,'\n',sep='')
  text=paste(text,'update ',as.integer(mcmc$nburn),"\n",sep='')
  text=paste(text,'coda *, stem("c:\\scripts\\graduacao\\modelo1\\CODA")\n
  exit\n',sep="")
  cat(text,file=file)
}

```

```
#####
# Código do arquivo comandos.cmd a ser executado pelo JAGS, gerado #
# automaticamente atraves do script jags.r #
#####

```

```
model in "c:\scripts\graduacao\modelo1\model.bug"
data in "c:\scripts\graduacao\modelo1\obitos-data.R"
compile, nchains(3)
parameters in "c:\scripts\graduacao\modelo1\obitos-par1.R", chain(1)
parameters in "c:\scripts\graduacao\modelo1\obitos-par2.R", chain(2)
parameters in "c:\scripts\graduacao\modelo1\obitos-par3.R", chain(3)
initialize
update 60000
monitor set q.est, thin(6)
update 120000
coda *, stem("c:\scripts\graduacao\modelo1\CODA")
exit

```



```
#####
# Codigo para implementacao do Modelo 1 (modelo1.txt) no JAGS          #
#  construido no diretorio "C:\scripts\graduacao\modelo1"             #
#####

### Funcao de verossimilhanca
model{
  for (i in 1:length(d)){
    d[i]~dpois(lambda[i])
    lambda[i]<-theta[i]*l[i]
  }

### Estabelecimento das prioris dos theta[i]
  for (i in 1:length(d)){
    theta[i]~dgamma(0.001,0.001)
  }
  theta[1:length(d)]<-sort(theta)

### Estabelecimento da distribuicao preditiva do numero de mortes
  for (i in 1:length(d)){
    drep[i]~dpois(lambda[i])
    theta.est[i]<-drep[i]/e[i]
    q.est[i]<-1-exp(-theta.est[i])
  }
}
```

## F) Modelo Estático Local para Construção da Tábua de Mortalidade

```
#####
# Código construído para executar o Modelo Estático Local no #
# WinBUGS a partir do R, através do pacote R2WinBUGS #
#####

### Carregamento do pacote R2WinBUGS e entrada de informações
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iterações:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento:"))
n.thin=as.numeric(readline("Informe o tamanho do intervalo (thin):"))
T=as.numeric(readline("Informe o número de anos observados:"))

### Especificação do diretório de trabalho e variável monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","estatloc2.txt")
WD="c:\\scripts\\graduacao\\estatloc2\\"

### Carga e criação do arquivo de dados
e=read.table(file=paste(WD,"expostos.txt", sep=""))
d=read.table(file=paste(WD,"obitos.txt", sep=""))
n=nrow(d)
e=t(e);d=t(d)
dim(e)<-c(T,n,1);dim(d)<-c(T,n,1)
data=list("n","d","e","T")

### Estabelecimento dos valores iniciais das cadeias
ini=read.table(file=paste(WD,"inicio_phi.txt", sep=""), header=T)
inits<-list(list(phi=ini$AT2000),list(phi=ini$AT83),
list(phi=ini$AT49))

### Estabelecimento de parâmetros para a função bugs e execução do
### WinBUGS, gerando traços das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=3, n.iter,
n.burnin, n.thin, bugs.directory="c:\\Program Files\\WinBUGS14\\",
program="WinBUGS", debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
  quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
  ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatísticas
### decorrentes da implementação do modelo e diagnósticos de convergência
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execução desta linha, codaPkg=F na função bugs
plot(out.sim) # para execução desta linha, codaPkg=F na função bugs
```

```
#####
# Codigo para implementacao do Modelo Estatico Local (estatloc.txt)      #
# no WinBUGS construido no diretorio:                                   #
# "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"         #
#####

### funcao de verossimilhanca

model{
  for (t in 1:T){
    for (i in 1:n){
      lambda[i,t]<-theta[i]*e[t,i,1]
      d[t,i,1]~dpois(lambda[i,t])
    }
  }

###Estabelecimento das prioris de phi[i,1]

  for (i in 1:n){
    phi[i]~dgamma(1.033,306.87)
  }

###Reparametrizacao

  theta[1]<-phi[1]

  for (i in 2:n){
    theta[i]<-theta[i-1]+phi[i]
  }

###Distribuicao Preditiva

  for (i in 1:n){
    for (t in 1:T){
      drep[i,t]~dpois(lambda[i,t])
    }
    theta.est[i]<-(drep[i,1]+drep[i,2]+drep[i,3]+drep[i,4]+drep[i,5])/
    (e[1,i,1]+e[2,i,1]+e[3,i,1]+e[4,i,1]+e[5,i,1])
    q.est[i]<-1-exp(-theta.est[i])
  }
}
```

### G) Modelo Estático Global para Construção da Tábua de Mortalidade

```
#####
# Código construído para executar o Modelo Estático Global no #
# WinBUGS a partir do R, através do pacote R2WinBUGS #
#####

### Carregamento do pacote R2WinBUGS e entrada de informações
require(R2WinBUGS)
n.iter=as.numeric(readline("Informe o total de iterações:"))
n.burnin=as.numeric(readline("Informe o tamanho do aquecimento (burn in)
:"))
n.thin=as.numeric(readline("Informe o tamanho do intervalo (thin):"))
T=as.numeric(readline("Informe o número de anos considerados:"))
b=as.numeric(readline("Informe o valor de beta (>0):"))
g=as.numeric(readline("Informe o valor de gama (>=1):"))

### Especificação do diretório de trabalho e variável monitorada
parameters<-c("q.est")
model.file<-system.file(package="R2WinBUGS","model","estatglob1.txt")
WD="c:\\scripts\\graduacao\\estatglob1\\"

### Carga e criação do arquivo de dados
e=read.table(file=paste(WD,"expostos.txt", sep=""))
d=read.table(file=paste(WD,"obitos.txt", sep=""))
n=nrow(d)
e=t(e);d=t(d)
dim(e)<-c(T,n,1);dim(d)<-c(T,n,1)
data=list("n","d","e","T")

### Estabelecimento dos valores iniciais dos parâmetros
inits1<-list(beta=b,gama=g)
inits<-list(inits1)

### Estabelecimento de parâmetros para a função bugs e execução do
### WinBUGS, gerando traços das cadeias e densidades a posteriori
out.sim=bugs(data, inits, parameters, model.file, n.chains=1, n.iter,
n.burnin, n.thin, bugs.directory="c:\\Program Files\\WinBUGS14\\",
program=c("WinBUGS"), debug=T, working.directory=WD, codaPkg=T)

for (i in 1:length(out.sim)){
  output=read.coda(out.sim[i],paste(WD,"codaindex.txt", sep=""),
quiet=F)
  plot(output, trace=T, density=T, smooth=T, auto.layout=T,
ask=dev.interactive())
}

### Retorna os valores das taxas graduadas, estimativas e estatísticas
### decorrentes da implementação do modelo e diagnósticos de convergência
logfile=bugs.log(paste(WD,"log.txt", sep=""))
print(logfile)
print(out.sim) # para execução desta linha, codaPkg=F na função bugs
plot(out.sim) # para execução desta linha, codaPkg=F na função bugs
```

```
#####
# Codigo para implementacao do Modelo Estatico Global (estatglob.txt)      #
# no WinBUGS construido no diretorio:                                     #
# "C:\Arquivos de programas\R\R-2.9.1\library\R2WinBUGS\model"          #
#####

### funcao de verossimilhanca

model{
  for (i in 1:n){
    for (t in 1:T){
      log(lambda[i,t]) <- log(e[t,i,1]) + log(beta) + (41+i) * log(gama)
      d[t,i,1] ~ dpois(lambda[i,t])
    }
    theta[i] <- beta * pow(gama, (41+i))
  }
}

###Estabelecimento das prioris

beta ~ dnorm(0,0.0001) I(0,)
gama ~ dnorm(0,0.0001) I(1,)

###Distribuicao Preditiva

for (i in 1:n){
  for (t in 1:T){
    drep[i,t] ~ dpois(lambda[i,t])
  }
  theta.est[i] <- (drep[i,1] + drep[i,2] + drep[i,3] + drep[i,4] + drep[i,5]) /
  (e[1,i,1] + e[2,i,1] + e[3,i,1] + e[4,i,1] + e[5,i,1])
  q.est[i] <- 1 - exp(-theta.est[i])
}
}
```

## H) Códigos Elaborados para Simulação por bootstrap e Intervalos de Confiança

```
#####
#Codigo construido no paoote R para informacao de parametros e      #
# carregamento dos arquivos necessarios a simulacao por bootstrap e  #
# construcao dos intervalos de confianca empiricos                   #
#####

n<-as.numeric(readline("Informe o numero de simulacoes que deseja gerar:"))
M<-as.numeric(readline("Informe a media do erro aleatorio da serie Kt:"))
SD<-as.numeric(readline("Informe o desvio padrao do erro aleatoria da serie
Kt:"))
MA<-as.numeric(readline("Informe o parametro media movel estimado:"))
C<-as.numeric(readline("Informe o termo constante estimado:"))
AJUSTE=read.table(file="c:\\Scripts\\Lee-Carter\\ajuste.txt",header=T)
DATA=read.table(file="c:\\Scripts\\Lee-Carter\\dados.txt",header=T)
source("c:\\Scripts\\Lee-Carter\\boot.r")
a<-bootstrap(n, M, SD, DATA, AJUSTE, MA, C)
```

```
#####
# Codigo construido no pacote R para gerar simulacoes dos valores de kt e #
# das taxas brutas de mortalidade, construindo dos intervalos de confianca#
# empiricos #
#####

bootstrap=function(n, M, SD, DATA, AJUSTE, MA, C){
  residuos<-matrix(rnorm(n*8,M,SD), nrow=n, ncol=8)
  K=matrix(nrow=n, ncol=8)
  Tx=matrix(nrow=49, ncol=n)
  require(graphics)
  T=2000 # primeiro ano calendario
  Min<-matrix(1,nrow=49, ncol=8)
  medias<-Max<-matrix(0,nrow=49, ncol=8)

  ### Monta os Kt's a partir da matriz de residuos
  for (i in 1:n){
    d=NULL
    for (j in 1:8){
      d[j]=C-(MA*AJUSTE[50+j,3])+residuos[i,j]
      if (AJUSTE[50+j,3]==0)
        d[j]=C-(MA*residuos[i,j-1])+residuos[i,j]
      K[i,j]=d[j]+AJUSTE[50+j,1]
      if (AJUSTE[50+j,1]==0) K[i,j]=d[j]+K[i,j-1]
    }
  }

  ### Monta a matriz de taxas brutas de mortalidade simuladas realizando n
  ### simulacoes para cada kt
  par(mfrow=c(3,3))
  for (J in 1:8){
    for (I in 1:n){
      for (l in 1:43){
        Tx[l,I]=exp(DATA[l,2]+DATA[l,3]*K[I,J])
        Min[l,J]<-min(Min[l,J],Tx[l,I])
        Max[l,J]<-max(Max[l,J],Tx[l,I])
      }
    }
    ts.plot(Tx, xlab="Idades", ylab="Taxas Brutas",
            main=paste("Replicacoes das Tx. Brutas - k",T+J))
  }

  ### Construcao dos Intervalos de confianca empiricos das taxas brutas
  ### de mortalidade
  leg=c("Limite inferior","Media","Limite Superior")
  medias[,1]=(Min[,1]+Max[,1])/2
  IC<-cbind(Min[,1],medias[,1],Max[,1])
  matplot(42:90,IC, type="l", lwd=1, lty=1, col=2:1, xlab="Idades",
          ylab="Taxas Brutas")
  legend("topleft", leg, col=2:1, text.col=1, lty=1, merge=T, lwd=1)
  for (i in 2:8){
    medias[,i]=(Min[,i]+Max[,i])/2
    ic<-cbind(Min[,i],medias[,i],Max[,i])
    IC<-cbind(IC,ic) # concatena ic a IC
    matplot(42:90,ic, type="l", lwd=1, lty=1, col=2:1, xlab="Idades",
            ylab="Taxas Brutas")
    legend("topleft", leg, col=2:1, text.col=1, lty=1, merge=T, lwd=1)
  }
  return(list(Min=Min,Max=Max,medias=medias))
}

```