Universidade Federal do Rio de Janeiro

Modelagem Bayesiana da Evolução da Mortalidade por Doenças Isquêmicas do Coração no Estado do Rio de Janeiro no Período de 1980 a 2002

Carla de Souza Lôbo

2007

Livros Grátis

http://www.livrosgratis.com.br

Milhares de livros grátis para download.

UFRJ

Modelagem Bayesiana da Evolução da Mortalidade por Doenças Isquêmicas do Coração no Estado do Rio de Janeiro no Período de 1980 a 2002

Carla de Souza Lôbo

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências Estatísticas.

Orientadores:
Helio S. Migon e Dani Gamerman

Rio de Janeiro
Maio de 2007

Modelagem Bayesiana da Evolução da Mortalidade por Doenças Isquêmicas do Coração no Estado do Rio de Janeiro no Período de 1980 a 2002

Carla de Souza Lôbo

Orientadores:

Helio S. Migon e Dani Gamerman

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários para obtenção do grau de Mestre em Ciências Estatísticas.

Aprovada por :

Presidente, Prof. Helio dos Santos Migon

Co-orientador, Prof. Dani Gamerman

Prof$^{\underline{a}}$ Flávia Maria Pinto Ferreira Landim

Prof. Nelson Albuquerque de Souza e Silva

Rio de Janeiro

Maio de 2007

Lôbo, Carla de Souza

Modelagem bayesiana da evolução da mortalidade por doenças isquêmicas do coração no estado do Rio de Janeiro no período de 1980 a 2002 / Carla de Souza Lôbo. - Rio de Janeiro: UFRJ / IM, 2007.

xi, 88 f. : il. ; 31cm.

Orientadores: Helio dos Santos Migon e Dani Gamerman

Dissertação (mestrado) - UFRJ / IM / Programa de Pós-Graduação em Estatística, 2007.

Bibliografia: f. 79-86.

1. Doenças Isquêmicas do Coração. 2. Força de Mortalidade. 3. Inferência Bayesiana. 4. Modelos Lineares Generalizados Dinâmicos. 5. Modelos Hierárquicos. I. Migon, Helio S. II. Gamerman, Dani. III. Universidade Federal do Rio de Janeiro, Instituto de Matemática. IV. Título.

Resumo

Modelagem Bayesiana da Evolução da Mortalidade por Doenças Isquêmicas do Coração no Estado do Rio de Janeiro no Período de 1980 a 2002

Carla de Souza Lôbo

Orientadores:

Helio S. Migon e Dani Gamerman

Atualmente, as doenças do aparelho circulatório são as principais causas de morte no Brasil, sendo as doenças isquêmicas do coração (DISC) uma das suas importantes subdivisões. Tem sido observado, nas últimas décadas, declínio da mortalidade por DISC. A modelagem do declínio destas taxas no estado do Rio de Janeiro, no período de 1980 a 2002, por sexo e faixa etária, é o objeto deste estudo. A análise desenvolvida baseou-se na utilização de um modelo Poisson para o número de mortos por DISC, para ambos os sexos, aplicando-se os conceitos de modelos hierárquicos e modelos lineares generalizados dinâmicos (West *et al.*, 1985) por meio de uma abordagem bayesiana. O conceito de fator de redução de mortalidade (Renshaw e Haberman, 2003) foi também aplicado ao conjunto de dados, avaliando-se o efeito da inclusão de pontos de quebra em alguns períodos da série, a fim de permitir uma maior flexibilidade na estrutura dos modelos. Para o ajuste foram utilizados algoritmos baseados no método de Monte Carlo via Cadeias de Markov (MCMC). Por fim, como critério de seleção de modelos, utilizou-se o *Deviance Information Criterion* - DIC (Spiegelhalter *et al.*, 2002).

Palavras-chave: doenças isquêmicas do coração; força de mortalidade; inferência bayesiana; modelos lineares generalizados dinâmicos; modelos hierárquicos.

Abstract

Bayesian Modeling of the Evolution of Mortality for Ischemic Heart Diseases in the State of Rio de Janeiro in the period of 1980 to 2002

Carla de Souza Lôbo

Supervisors:

Helio S. Migon and Dani Gamerman

Currently, the diseases of the circulatory system are the main causes of deaths in Brazil, and ischemic heart disease (DISC) one of its major subdivisions. It has been observed in recent decades, declining mortality DISC. The modeling of the decline of these rates in the state of Rio de Janeiro, in the period from 1980 to 2002, by sex and age group, is the object of this study. The analysis developed was based on the use of a Poisson model for the number of deaths by DISC, for both sexes, where we applied the concepts of hierarchical models and generalized linear dynamic models (West *et al.*, 1985) through a bayesian approach. The concept of a mortality reduction factor (Renshaw and Haberman, 2003) was also applied to the collection of data, where the effect of including items of breakage in some periods of the series was evaluated in order to enable greater flexibility in the structure of the models. For the adjustment, Markov Chains Monte Carlo (MCMC) methods were used. Finally, as a criterion for model selection, it was used the Deviance Information Criterion - DIC (Spiegelhalter *et al.*, 2002).

Key-words: ischemic heart diseases; force of mortality; bayesian inference; generalized linear dynamic models; hierarchical models.

Agradecimentos

A Deus, pois foi da vontade Dele a conclusão deste curso. Agradeço por ter me capacitado e por ter me dado ânimo. Porque Dele, e por Ele, e para Ele são todas as coisas; glória, pois, a Ele eternamente. Amém! (Rm 11, 36).

Ao Gustavo - que foi promovido de namorado a noivo, e depois de noivo a marido durante o curso de mestrado - por ser uma pessoa fantástica! Agradeço por estar sempre ao meu lado e pela ajuda em todos os momentos.

Aos meus pais, Luiz Carlos e Ana Maria, e à minha irmã Erika, pelo apoio, compreensão e torcida em mais esta etapa da minha vida. Amo vocês!

À minha cadela-irmã Shasnay, pelos momentos de alegria, amor e cumplicidade.

Aos meus sogros, Glênio e Graziela, e minha cunhada Gisele, que sempre oraram e torceram pela conclusão deste trabalho.

Aos amigos da minha turma de mestrado - Valmária, Carol, Ana Paula, Lúcia e Marcus Vinícius - e para todos os outros amigos feitos no DME. Agradeço pela amizade e pelos momentos divertidos!

Aos amigos do Miolo - Cristiane, Aline, Leo e Thaís - pelas risadas e pelos Miolos. Em especial ao casal Leo e Thaís pela ajuda com o WinEdt e com o Abstract, respectivamente.

Ao Oswaldo, que não apagava o *scrap* com o prazo da defesa.

Aos amigos Aline, Waldes e Carla pela motivação a continuar neste desafio.

A todos que me deram força e não me deixaram desistir.

À Igreja Batista do Meier pelas orações.

Aos professores Helio Migon e Dani Gamerman pela orientação, compreensão e paciência, e também a todos os outros professores do DME/UFRJ pela transmissão de conhecimento.

Ao Eduardo Thompson (Secretaria de Pós-Graduação do IM/UFRJ) pela ajuda na parte burocrática.

Ao CNPQ pelo apoio financeiro.

Agradeço pelo fornecimento dos dados: Dr. Bernardo Tura, Dra. Gláucia Maria Moraes de Oliveira e Dra. Ângela Cascão da Secretaria de Estado de Saúde do Rio de Janeiro.

Sumário

1 Introdução **1**

1.1 Doenças Isquêmicas do Coração 1

1.2 Redução da Mortalidade por DISC 3

1.3 Fatores de Influência 6

1.4 Importância do estudo da evolução das DISC 9

1.5 Modelagem da Mortalidade 10

1.6 Objetivos do Estudo 11

2 Conceitos Básicos sobre Mortalidade **13**

2.1 Força de Mortalidade 13

2.2 Lei de Mortalidade de Gompertz 14

3 Inferência Bayesiana **16**

3.1 Teorema de Bayes 16

3.2 Distribuição Preditiva 18

3.3 Estimação de Parâmetros 18

3.4 Inferência via Simulação Estocástica 20

3.4.1 Inferência via MCMC 21

3.4.2 WinBUGS 24

4 Modelos Lineares **26**

4.1 Modelos Lineares Generalizados 27

4.2 Modelos Lineares Dinâmicos 29

4.3 Modelos Lineares Generalizados Dinâmicos 31

4.4 Modelos Lineares Hierárquicos 32

5 Modelos Propostos **33**

5.1 Modelagem do número de óbitos por DISC 33

5.2 Modelo Dinâmico Local 34

5.3 Modelo Dinâmico Global 35

5.4 Modelo Hierárquico 36

5.5 Modelo Linear Generalizado com Fator de Redução de Mortalidade 37

5.5.1 Fator de Redução de Mortalidade 37

5.5.2 O Modelo 38

6 Resultados **40**

6.1 Descrição dos Dados 40

6.2 Modelo Dinâmico Local 44

6.3 Modelo Dinâmico Global 51

6.4 Modelos Hierárquicos 58

6.5 Modelo Linear Generalizado com Fator de Redução de Mortalidade 65

7 Considerações Finais **77**

Referências Bibliográficas **79**

Apêndice **87**

Capítulo 1

Introdução

1.1 Doenças Isquêmicas do Coração

A partir da segunda metade do século XX, na maioria dos países desenvolvidos e também naqueles em desenvolvimento, as Doenças Isquêmicas do Coração (DISC) vêm se destacando como uma das principais causas de morte, com taxas de mortalidade de elevada magnitude (Moraes *et al.*, 2000).

No Brasil, as doenças do aparelho circulatório (DAC) constituem a principal causa de morte no país (Daniel et al, 2005). Com relação a estas doenças, para o total dos óbitos e para ambos os sexos, predominam as DISC, atingindo mais os homens, e as Doenças Cerebrovasculares, atingindo mais as mulheres (Tomassini *et al.*, 2003).

Verifica-se também que, nos anos mais recentes, as DISC já predominam não só em São Paulo como também no Rio Grande do Sul, pelo menos entre homens de 30 a 79 anos (Oliveira *et al.*, 2005 [b]).

Na cidade de Curitiba, em 1998, as DAC representaram 33.8% do total de óbitos, sendo as DISC e, principalmente, o infarto agudo do miocárdio o seu maior componente (Daniel *et al.*, 2005).

Segundo dados da Prefeitura Municipal da Cidade do Rio de Janeiro, referentes ao período de 1997 a 2003, as DISC são a principal causa de óbitos para ambos os sexos a partir dos cinqüenta e cinco anos de idade. Para a faixa etária de 45 a 54, as DISC são a segunda maior causa de óbitos, atrás apenas daquelas causadas por Doenças Mal Definidas. Adicionalmente, elas representam a quarta maior causa de morte para pessoas com idades entre 35 e 44 anos (Fonte: Armazém de Dados - Prefeitura Municipal do Rio de Janeiro).

Um maior número de casos existentes de DISC em idosos é um fenômeno esperado e observado em estudos, como em Laurenti *et al.* (1981), uma vez que este tipo de doença possui a aterosclerose como causa principal. Esta última possui progressão lenta, podendo ter início na infância e culminando com o aparecimento de síndromes isquêmicas sintomáticas na vida adulta e, principalmente, em idosos (*Inquérito domiciliar sobre comportamentos de risco e morbidade referida de doenças e agravos não transmissíveis: Brasil, 15 Capitais e Distrito Federal 2002-2003*).

A morbidade por DISC representa uma grande carga para o país. Estudos indicam que a eliminação total dos grupos de causas como fatores de risco de morte por doenças do aparelho circulatório pode propiciar ganhos na esperança de vida, para homens e mulheres, respectivamente de 10,4 e 10,8 anos. Este número é bem superior ao ganho obtido com a eliminação dos grupos de causas associados às doenças infecciosas e parasitárias (7,9 e 8,1 anos), neoplasmas malignos (6,0 e 6,6 anos), doenças do aparelho respiratório

(6,5 e 6,7 anos) e causas externas (7,0 e 5,2 anos) (Paes, 1985).

1.2 Redução da Mortalidade por DISC

Apesar de apresentarem uma taxa de mortalidade muito superior à do Brasil - em parte devido à juventude da população deste último (Laurenti, 1982) - a tendência dessas taxas tem declinado nos últimos 40 anos em países como Estados Unidos, Canadá, Austrália, Japão, Reino Unido e outros da Europa Ocidental, após um período praticamente estacionário. No Brasil, a partir dos anos 70, alguns estudos têm mostrado a tendência de declínio da mortalidade por DISC, como Daniel *et al.* (2005) e Mansur *et al.* (2001).

Em estudo do declínio das taxas de mortalidade do período de 1950 a 1994 nos países que compõem o G7 - Canadá, França, Alemanha, Itália, Japão, Reino Unido e EUA -, argumentou-se que tal fato pode ser resultado de uma aplicação de recursos e conhecimentos em saúde pública e redução de mortalidade. Esta aplicação de recursos poderia produzir mudanças exponenciais nas taxas de mortalidade, embora não de forma constante. O balanço entre os recursos empregados e a sua efetividade afetaria o declínio da mortalidade ao longo dos anos (Tuljapurkar *et al.*, 2000).

No Brasil, quando observamos o perfil da mortalidade proporcional por grandes grupos de causas, verificamos que esta vem se modificando gradualmente desde a década de 1940 até os dias atuais. Observamos uma redução progressiva da importância relativa das doenças infecciosas e parasitárias e um aumento concomitante das DAC e das doenças neoplásicas. Esta redução, associada com a diminuição das taxas de mortalidade infantil, produziu um

aumento na expectativa de vida e, ainda mais, começou-se a observar uma estabilização e posterior quedas graduais nas taxas de mortalidade cardiovascular, principalmente a partir de 1980 (Oliveira *et al.*, 2003).

Entre 1980 e 2000, uma tendência de redução na mortalidade por DISC já pôde ser observada em algumas regiões, evidenciando-se uma tendência de redução nas regiões Sul e Sudeste, conforme consta do *Inquérito domiciliar sobre comportamentos de risco e morbidade referida de doenças e agravos não transmissíveis: Brasil, 15 Capitais e Distrito Federal 2002-2003* e em Oliveira *et al.*, (2006). No mesmo sentido, evidenciou-se uma redução das internações por DISC de 29,8% nos hospitais do SUS, no período de 1993 a 1997 (Laurenti *et al.*, 2000).

Com exceção de um pequeno aumento observado em algumas capitais, para algumas determinadas faixas etárias, foi detectado um declínio nas taxas de mortalidade por DISC na maioria das capitais brasileiras no período de 1980 a 1998 (Oliveira *et al.*, 2005 [a]).

Em estudo realizado na cidade de Curitiba, os resultados permitiram identificar uma redução nos coeficientes de mortalidade, ajustados por idade, no infarto agudo do miocárdio e no restante das doenças isquêmicas, mostrando uma redução das DISC de 1980 a 1998, em ambos os sexos. Esta redução em ambos os sexos, pode ser, em parte, explicada por uma melhor contribuição da assistência médica, bem como das medidas de prevenção primária para estas doenças (Daniel *et al.*, 2005). Os mesmos resultados e conclusões foram obtidos analisando-se a série histórica das DISC em Goiânia, no período de 1980 a 1994 (Moraes *et al.*, 2000).

Apesar do declínio observado em várias capitais no Brasil, observa-se

um nítido efeito de período no ano de 2000 no Estado do RJ. Supõe-se que este fato seja reflexo de um efeito próprio da década de 1990, uma vez que neste período não foi encontrado relato de implementação de melhorias na assistência médica ou no controle efetivo dos fatores de risco na população do Estado do RJ que tivessem impacto significativo na modificação das taxas de mortalidade por DISC (Oliveira *et al.*, 2005 [a]).

Ainda como hipóteses para explicar este declínio, pode-se pensar na efetiva diminuição dos óbitos em virtude da crescente mortalidade por outras causas, como por exemplo a AIDS e causas externas, ou ainda pelo aumento da oferta de tecnologias específicas para as doenças desta natureza. Nesta linha, de acordo com Tomassini *et al.* (2003) e Passos *et al.* (2000), não seria possível inferir que a morbidade esteja diminuindo, pois sua letalidade poderia estar sendo atenuada e a mortalidade desviada para outras causas.

Assim, a ausência de estudos para os fatores de risco ou de proteção ou para a incidência desta doença, assim como de estudos analíticos que possam testar sua associação entre a mortalidade, limitam as explicações desejadas para o declínio observado (Daniel *et al.*, 2005).

Analisando a redução da mortalidade por DISC por sexo, conclui-se que esta vem ocorrendo de forma diferenciada, dependendo da região e do período considerado. Mansur *et al.* (2001) observaram uma redução significativamente mais pronunciada entre os homens do que nas mulheres no período 1979 a 1996. Conclusão semelhante foi obtida em estudo realizado em Salvador (Passos *et al.*, 2000). Por outro lado, em estudo realizado na cidade de São Paulo de 1950 a 1981, observou-se uma redução mais acentuada para as mulheres (Lolio e Laurenti, 1986). Este último, por sua vez, não analisa o período em que a redução foi mais acentuada para ambos os sexos,

ocorrida nas duas décadas seguintes.

1.3 Fatores de Influência

Os desfechos clínicos ou mortes por DISC são decorrentes de complexas interações entre as populações e o ecossistema em que vivem, desde o seu nascimento. Os estudos que investigam as causas de morte das populações precisam ser analisados dentro desse contexto ambiental. Assim, o evento que causou a morte pode ser decorrente de nosso limite biológico de vida, para o qual não se conhecem os determinantes, como também pela interação entre as populações e os ambientes nos quais as mesmas estejam inseridas, desde a vida intra-uterina até a morte. A complexa interação entre o conjunto de genes de cada indivíduo e o modo como funcionam ou se inter-relacionam com os múltiplos fatores ambientais irá determinar as condições de saúde e morte das populações. Assim, a morte nas DISC é precedida pela presença de fatores de risco, alguns já conhecidos (hipertensão arterial, diabetes, tabagismo, história familiar, etc) e outros ainda prováveis (processos inflamatórios ou infecciosos crônicos) ou desconhecidos, que se relacionam com comportamentos individuais ou familiares, como alcoolismo, hábitos alimentares, sedentarismo, condições de vida, de habitação e de trabalho. A ocorrência destes fatores e sua prevalência na população é dependente do modo como a sociedade se organiza para promover a saúde, ou seja, da atuação dos serviços de saúde, das políticas governamentais e intersetoriais (Oliveira *et al.*, 2005 [a]).

Assim, em vários estudos tais como em Lotufo e Lolio (1993) e Mansur *et al.* (2001), procurou-se atribuir a redução da mortalidade por DISC ao

controle dos fatores de risco cardiovascular clássicos.

Entretanto, os controles dos fatores de risco, mesmo em países desenvolvidos, ainda não atingiram os níveis adequados e, além do mais, seria necessário que o controle desses fatores de risco se fizesse antes do reconhecimento da queda da mortalidade para que se pudesse atribuir ao mesmo este efeito.

Oliveira *et al.* (2003), em estudo crítico das mudanças das taxas de mortalidade ocorridas entre 1980 e 2000, no estado do Rio de Janeiro, sugerem que o desenvolvimento econômico ocorrido nas décadas anteriores a 1980, mais do que o duvidoso controle dos fatores de risco cardiovascular clássicos, teria contribuído para a queda das taxas de mortalidade. Assim, os autores ressaltam a necessidade de sugerir outras hipóteses para explicar a tendência desta redução, mesmo porque fatores biomédicos e de estilo de vida explicam apenas pequena proporção da variabilidade das doenças cardiovasculares. Neste sentido, os autores apontam para a adoção de estratégias que venham a melhorar as condições de vida da população, a fim de reduzir as enormes desigualdades sociais, as quais podem ter maior impacto na redução da mortalidade do que o controle dos fatores de risco clássicos, que por sua vez, devem continuar. Conclusões semelhantes foram observadas em Oliveira *et al.* (2005 [d]), Oliveira *et al.* (2006), Tura *et al.* (2006) e Moraes *et al.* (2000).

Uma menor prevalência de DISC foi observada nos entrevistados com escolaridade mais alta em sete capitais do Brasil. Entretanto, estes dados não podem ser interpretados com uma associação entre escolaridade e prevalência dessa patologia, já que estas doenças ocorrem em coortes mais idosas, que têm menor escolaridade (*Inquérito domiciliar sobre comportamentos de risco*

e morbidade referida de doenças e agravos não transmissíveis: Brasil, 15 Capitais e Distrito Federal 2002-2003).

Marcopito (2003), em estudo realizado no nordeste, avaliou que a situação de aumento da mortalidade por DISC em comparação às demais causas não parece dever-se à falta de acesso à assistência hospitalar, já que os brasileiros da região nordeste exibiram a maior percentagem de óbitos por DISC ocorridos em hospital em todos os anos do período estudado. Os autores sugerem que as causas devem ter origem mais remota e estar ligadas às condições de vida e alimentação, à atenção primária e secundária à saúde.

Ainda outros estudos podem ser citados associando aspectos econômicos e sociais com as taxas de mortalidade por DISC. Neste sentido, uma relação entre o PIB do Brasil, defasado em 20 anos, e a mortalidade por DISC foi evidenciado por Tura *et al.* (2006).

Analisando diversos estudos já realizados no Brasil, ainda não se conhece, com precisão, a participação dos fatores envolvidos na determinação do declínio da mortalidade pelas doenças isquêmicas do coração.

Neste sentido, um estudo realizado no Rio de Janeiro concluiu que são necessárias pesquisas correlacionando as taxas de mortalidade com indicadores socioeconômicos de cada região, bem como providenciar a implantação de um Serviço de Verificação de Óbitos para se conhecer melhor as tendências reais da mortalidade por DISC (Oliveira *et al.*, 2005 [c]). A dificuldade de se obter registros de óbitos confiáveis é considerada um problema mundial (Mansur *et al.*, 2001).

Além das dificuldades já citadas, o problema das imprecisões nas declarações da causa da morte cria grandes barreiras que limitam as conclusões

dos estudos a respeito de mortalidade por doenças, tais como as DISC. Estas imprecisões condicionam o aumento da proporção de causas mal definidas, comprometendo a qualidade do indicador. Isto pode produzir impacto imediato no estudo das DISC, uma vez que o seu diagnóstico depende de condições assistenciais adequadas (fonte: DATASUS). Oliveira *et al.* (2005) [d] destaca um aumento da mortalidade por causas mal definidas no estado do Rio de Janeiro a partir de 1990, supondo que o mesmo está relacionado à portaria de 29 de janeiro de 1990. Esta portaria determinou a declaração nos atestados médicos de morte por *causa indeterminada* assim que forem esgotadas todas as tentativas de se determinar a causa básica do óbito, descartada suspeita de causa violenta (acidente, homicídio ou suicídio). Por este motivo, alguns estudos, como o citado, trabalharam com taxas de mortalidade compensadas. Esta abordagem, porém, não está no presente escopo.

1.4 Importância do estudo da evolução das DISC

A análise das variações geográficas e temporais da mortalidade específica por DISC, identificando situações que possam demandar a realização de estudos especiais, podem subsidiar processos de planejamento, gestão e avaliação de políticas e ações preventivas e assistenciais relativas às DAC (Paes, 1985).

É importante não esquecer que, mesmo com o declínio das taxas de mortalidade por DISC, as mesmas permanecem como principal causa de óbito, merecendo atenção constante e prioritária pelos profissionais de saúde e serviços públicos na prevenção primária e secundária (Daniel *et al.*, 2005).

1.5 Modelagem da Mortalidade

O surgimento da causa ou causas de morte no homem se dá pela sua exposição permanente ao risco de contrair uma ou mais enfermidades ao longo de sua vida. Os dados que representam estes riscos podem ser obtidos por meio da análise da evolução de tábuas de mortalidade associadas a algum fator de risco.

Para esta análise, geralmente são formuladas leis de mortalidade que fornecem uma medida dos riscos de morte associados à alguma causa específica em uma determinada época para uma determinada faixa etária.

Na análise de tábuas de mortalidade, vários modelos têm sido empregados. Entre eles, destacamos aqueles baseados na clássica lei de mortalidade de Gompertz, a qual sugere que após um certo estágio de maturidade sexual, o risco de morte passa a se comportar de forma exponencial (Gompertz, 1825).

Entre outros estudos realizados no Brasil, podemos destacar ainda a análise conduzida por Paes (1985), a qual, utilizando Tábuas de Vida de Múltiplo Decremento para a população do município de Recife em 1979, permitiu estimar o efeito produzido na mortalidade depois da eliminação de certos grupos de causas de morte, entre elas as DISC.

Entretanto, muitos estudos epidemiológicos a respeito da evolução das taxas de mortalidade baseiam-se no simples cálculo de coeficientes de correlação e em ajustes de modelos clássicos de regressão linear. Entre os estudos que utilizam esta modelagem na análise da evolução das taxas de mortalidade por DISC no Brasil, podemos citar Daniel *et al.* (2005), Oliveira *et al.*

(2005 [d]), Oliveira *et al.* (2006), Tura *et al.* (2006), Lolio *et al.* (1994), Moraes *et al.* (2000) e Mansur *et al.* (2001). Entretanto, a simples aplicação destes modelos não permite ao pesquisador captar a dependência temporal existente entre as taxas de mortalidade para uma determinada população.

1.6 Objetivos do Estudo

Neste trabalho serão utilizados modelos baseados em leis de mortalidade a fim de se estudar a evolução das tábuas de mortalidade por DISC no estado do Rio de Janeiro, no período de 1980 a 2002, avaliando-se os parâmetros associados de forma dinâmica no tempo sob uma abordagem bayesiana.

Para isto, o trabalho estruturou-se de tal maneira que no capítulo 2 são apresentados os conceitos básicos associados às leis de mortalidade e à lei de mortalidade de Gompertz.

Nos capítulos 3 e 4 são apresentados os principais aspectos da inferência bayesiana e dos modelos lineares, os quais constituem-se como base metodológica para análise dos dados utilizados neste trabalho.

No capítulo 5 são apresentados os diferentes modelos que foram propostos para a análise dos dados do estudo.

No capítulo 6 são apresentados e comparados os principais resultados da aplicação dos diferentes modelos propostos no capítulo 5 aos dados de mortalidade por DISC no estado do Rio de Janeiro.

Por fim, no capítulo 7 são apresentadas as principais conclusões obtidas com a aplicação dos diferentes modelos e são destacadas as contribuições

geradas para o estudo da evolução da mortalidade por DISC.

Capítulo 2

Conceitos Básicos sobre Mortalidade

Neste capítulo serão descritos alguns conceitos básicos associados às leis de mortalidade e à lei de mortalidade de Gompertz.

2.1 Força de Mortalidade

Seja X uma variável aletória contínua que representa a idade de morte de um recém nascido com função de distribuição denotada por $F(x)$. Tem-se que:

$$F(x) = P(X \leq x), \quad x \geq 0. \tag{2.1}$$

Define-se a função densidade de probabilidade de X por $f(x) = F'(x)$, isto é,

$$f(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}. \tag{2.2}$$

Considere a probabilidade condicional $P(x \leq X \leq x + \Delta x | X > x)$, isto é, a probabilidade de que um indivíduo venha a falecer durante as próximas Δx unidades de tempo, desde que tenha sobrevivido ao instante x. Então, com base em (2.2), segue que:

$$P(x \leq X \leq x + \Delta x | X > x) = \frac{F(x + \Delta x) - F(x)}{1 - F(x)} \cong \frac{\Delta x . f(x)}{1 - F(x)}, \tag{2.3}$$

supondo Δx pequeno.

Definimos, então, a *Força de Mortalidade* μ_x, uma medida de variação (ou taxa) instantânea da intensidade da morte, por:

$$\mu_x = \frac{f(x)}{1 - F(x)}, \quad \mu_x > 0. \tag{2.4}$$

2.2 Lei de Mortalidade de Gompertz

São chamados Leis de Mortalidade os modelos estocásticos que têm por objetivo estimar a mortalidade, a sobrevida ou o risco de morte em uma determinada idade.

Em 1825, o matemático inglês Benjamin Gompertz propôs um modelo paramétrico para a relação entre a força de mortalidade e a idade, após observar que a taxa de mortalidade humana aumenta exponencialmente com a idade depois da maturidade sexual. Este modelo é chamado de Lei de Gompertz e é definido a seguir:

$$\mu_x = B.c^x \tag{2.5}$$

sujeito às seguintes restrições: $B > 0$, $c \geq 1$ e $x \geq 0$, sendo x a idade em anos.

Na escala logarítmica, a equação anterior é:

$$log(\mu_x) = \log(B) + x.\log(c). \tag{2.6}$$

Seja $\alpha = \log(B)$ e $\beta = \log(c)$. A equação (2.6) será reescrita como:

$$log(\mu_x) = \alpha + \beta.x \tag{2.7}$$

satisfazendo $\alpha \in \Re$ e $\beta \geq 0$.

Tura (2006) interpretou biologicamente os parâmetros especificando a dependência com a idade e/ou o ambiente, conforme pode ser visto na Tabela 2.1.

Tabela 2.1: Parâmetros do modelo, sua definição, interpretação e relação com idade e ambiente.

Parâmetro	Definição	Representação	Relação com a idade	Relação com o ambiente
$B = \exp(\alpha)$	Hostilidade do meio ambiente	Risco ambiental e riscos competitivos	Independente	Dependente
$\log(c) = \beta$	Taxa exponencial de mortalidade	Risco ambiental	Dependente	Dependente

Capítulo 3

Inferência Bayesiana

Neste capítulo serão descritos os principais conceitos para realização de inferência estatística sob o ponto de vista bayesiano. Ao inferir a respeito de uma quantidade de interesse desconhecida e não observável θ, obtém-se conclusões baseadas em especificacões probabilísticas condicionadas à uma amostra de valores observados relacionados ao problema de interesse.

3.1 Teorema de Bayes

Todos os procedimentos de inferência, na abordagem bayesiana, têm como base o *Teorema de Bayes*, como descrito a seguir (Migon e Gamerman, 1999):

Seja θ uma quantidade de interesse desconhecida e não observável contida no conjunto de possíveis valores Θ, θ sendo escalar, vetor ou matriz. Seja $\mathbf{H}$ a informação inicial disponível para o parâmetro de interesse sumarizada probabilisticamente por $p(\theta|\mathbf{H})$. Se $\mathbf{H}$ é suficientemente informativa, temos

a descrição completa da incerteza a respeito de θ.

Se $\mathbf{H}$ não é informativa o bastante, devemos atualizar a informação a respeito de θ, observando-se uma amostra $\mathbf{x}$ de um vetor aleatório $\mathbf{X}$ relacionado com θ. Assim, a informação disponível para a inferência passará a ser $\mathbf{H}^* = \mathbf{H} \cap \{\mathbf{X} = \mathbf{x}\}$.

Para realizar a atualização é necessário conhecer a distribuição amostral de $\mathbf{X}$, conhecida como *função de verossimilhança*. Esta função associa a cada θ o valor $p(\mathbf{x}|\theta)$, isto é, fornece as chances de cada valor de θ ter levado àquele valor observado de $\mathbf{x}$.

Assim, temos:

$$p(\theta|\mathbf{H}^*) = p(\theta|\mathbf{x}, \mathbf{H}) = \frac{p(\theta, \mathbf{x}|\mathbf{H})}{p(\mathbf{x}|\mathbf{H})} = \frac{p(\mathbf{x}|\theta, \mathbf{H})p(\theta|\mathbf{H})}{p(\mathbf{x}|\mathbf{H})} \tag{3.1}$$

onde

$$p(\mathbf{x}|\mathbf{H}) = \int_{\Theta} p(\theta, \mathbf{x}|\mathbf{H}) d\theta.$$

Nota-se que $p(\mathbf{x}|\mathbf{H})$ não depende de θ e que $\mathbf{H}$ é comum a todos os termos. O teorema pode, então, ser reescrito da seguinte forma:

$$p(\theta|\mathbf{x}) \propto p(\mathbf{x}|\theta)p(\theta). \tag{3.2}$$

O resultado apresentado em (3.2) fornece a regra de atualização de probabilidades sobre θ, partindo de $p(\theta)$ e chegando a $p(\theta|\mathbf{x})$. Estas distribuições são conhecidas como distribuições *a priori* e *a posteriori*, respectivamente.

3.2 Distribuição Preditiva

Para calcular a constante removida do resultado apresentado em (3.2), basta reescrevê-lo como (Migon e Gamerman, 1999)

$$p(\theta|\mathbf{x}) = kp(\mathbf{x}|\theta)p(\theta)$$

onde

$$\begin{aligned} k^{-1} = p(\mathbf{x}|\mathbf{H}) &= \int_{\Theta} p(\mathbf{x}|\theta)p(\theta)d\theta \\ &= E_{\theta}[p(\mathbf{x}|\theta)]. \end{aligned}$$

Denotando $p(\mathbf{x}|\mathbf{H})$ como $p(\mathbf{x})$, temos que $p(\mathbf{x})$ é a *distribuição preditiva a priori de* $\mathbf{X}$ e, para um dado $\mathbf{H}$, comporta-se como uma predição.

Uma importante aplicação é a previsão de uma observação futura y dado que foi observado um conjunto de dados $\mathbf{y}$. Essa previsão baseia-se na distribuição de probabilidade de $Y|\mathbf{Y} = \mathbf{y}$, e, sabendo que Y e $\mathbf{Y}$ são independentes condicionalmente a θ, temos:

$$p(y|\mathbf{y}) = \int_{\Theta} p(y, \theta|\mathbf{y})d\theta = \int_{\Theta} p(y|\theta, \mathbf{y})p(\theta|\mathbf{y})d\theta = \int_{\Theta} p(y|\theta)p(\theta|\mathbf{y})d\theta. \quad (3.3)$$

Nota-se que a *distribuição preditiva a posteriori* de $Y|\mathbf{Y} = \mathbf{y}$ representa a verossimilhança esperada com relação a distribuição a posteriori de θ, ou seja,

$$p(y|\mathbf{y}) = E_{\theta|\mathbf{y}}[p(y|\theta)].$$

3.3 Estimação de Parâmetros

A estimação de parâmetros pode ser realizada de forma pontual ou intervalar.

(i) Estimação Pontual

Na inferência bayesiana, a estimação pontual de um parâmetro θ pode ser vista como um problema de decisão. Este problema de decisão está especificado quando são descritos os seguintes espaços (Migon e Gamerman, 1999):

- espaço do parâmetro Θ;
- espaço dos possíveis resultados do experimento Ω;
- espaço das possíveis ações A.

Uma *regra de decisão* δ é uma função definida em Ω com valores em A, isto é, $\delta : \Omega \to A$.

Uma *função perda* L é associada a cada regra de decisão $\delta(x)$, $x \in \Omega$, e a cada valor possível de $\theta \in \Theta$. Têm-se, então, uma medida de punição por tomar-se a decisão $\delta(x)$ quando o valor verdadeiro do parâmetro é θ.

As funções perdas mais comumente utilizadas são as simétricas, da forma $L(\delta, \theta) = h(\delta - \theta)$, para alguma função h.

A *perda esperada a posteriori* é o risco associado a uma determinada regra de decisão $\delta(x)$ e é definido como

$$R(\delta) = E_{\theta|x}(L(\delta, \theta)).$$

O estimador pontual de θ é a regra de decisão ótima, isto é, a que minimiza o risco esperado segundo uma função perda especificada.

A média, a mediana e a moda da distribuição a posteriori de θ são os estimadores pontuais associados às perdas quadráticas, perdas absolutas e perdas zero-um, respectivamente.

(ii) Estimação por Intervalo

Sob o ponto de vista bayesiano, a forma mais adequada de avaliar a informação disponível a respeito de parâmetros desconhecidos é por meio da distribuição a posteriori (Migon e Gamerman, 1999). Ao resumir a distribuição a posteriori em um único valor, não se tem uma medida da precisão da estimativa obtida. Uma alternativa para contornar esta situação é obter intervalos de credibilidade para estes valores.

A definição de intervalo de credibilidade bayesiano é descrita a seguir:

- seja θ uma quantidade desconhecida definida em Θ. Uma região $\mathbf{C} \subset \Theta$ é um *intervalo de credibilidade bayesiano* $100(1-\alpha)\%$ para θ se $p(\theta \in \mathbf{C}|\mathbf{x}) \geq 1 - \alpha$.

Neste caso, $1 - \alpha$ é chamado de *nível de credibilidade*, e o intervalo de credibilidade bayesiano com este nível é denotado por $IC(1 - \alpha)\%$.

3.4 Inferência via Simulação Estocástica

Na inferência bayesiana, a distribuição a posteriori é a descrição probabilística da informação disponível de uma quantidade de interesse (Migon e Gamerman, 1999). Quando a função densidade de probabilidade dessa distribuição não possui forma conhecida, ou é demasiadamente complicada, métodos computacionais intensivos para aproximação podem ser usados.

Simulações estatísticas são usadas para estimar características probabilísticas de modelos que não podem ser computados analiticamente (DeGroot, 2002). Sob o ponto de vista bayesiano, o interesse é obter amostras de distribuições a posteriori, para obter informações a seu respeito.

Ao lidar-se com problemas altamente complexos, devido ao grande número de parâmetros envolvidos, em geral utilizam-se métodos de *Monte Carlo via Cadeias de Markov* (MCMC) para realizar-se o processo inferencial.

3.4.1 Inferência via MCMC

A idéia básica dos métodos MCMC é construir uma cadeia de Markov cuja distribuição de equilíbrio é igual à distribuição de interesse. Após um número finito de simulações desta cadeia, espera-se atingir a distribuição de equilíbrio e, a partir deste ponto, dar origem a uma amostra da distribuição de interesse.

Considere que $\theta_1, ..., \theta_p$ possuem densidade conjunta $p(\boldsymbol{\theta}) = p(\theta_1, ..., \theta_p)$ e que $q(\theta, \theta^*)$ define a distribuição condicional das transições do estado θ.

É possível, então, construir uma cadeia com probabilidades de transição invariantes no tempo, onde cada estado pode ser obtido de um outro estado com um número finito de iterações. Assim, independemente do estágio inicial, uma trajetória pode ser gerada e, consequentemente, pode-se alcançar a distribuição de equilíbrio $p(\boldsymbol{\theta})$.

Existem muitos modos de se construir a cadeia de Markov. Dentre os métodos mais utilizados, temos o *amostrador de Gibbs*, proposto por Geman e Geman (1984) e popularizado por Gelfand e Smith (1990), e o *método de Metropolis-Hastings*, proposto por Metropolis *et al.* (1953) e Hastings (1970). Gamerman (1997) e Migon e Gamerman (1999) apresentam uma descrição detalhada desses dois métodos de simulação.

(i) Amostrador de Gibbs

O *amostrador de Gibbs* é um esquema iterativo de amostragem de uma

cadeia de Markov, cujas probabilidades de transição dos estados são realizadas a partir das distribuições condicionais completas.

Seja $\pi(\theta)$ a distribuição de interesse, onde $\theta = (\theta_1, ..., \theta_p)$. Deseja-se gerar uma amostra de π, e supõe-se que esta geração é complicada.

Considere $\pi_i(\theta_i) = p_i(\theta_i|\boldsymbol{\theta}_{-i})$ como a função densidade condicional de θ_i, dados os valores de todos os outros θ_j's $(j \neq i)$, e assume-se que é possível amostrar valores destas distribuições para cada $i = 1, ..., p$.

Especifica-se, então, valores iniciais $\boldsymbol{\theta}^0 = (\theta_1^0, ..., \theta_p^0)$ para todas as quantidades desconhecidas do modelo. Assim, na j-ésima iteração a cadeia se encontra no estado $\boldsymbol{\theta}^{(j)}$ e a posição da cadeia na iteração seguinte $(j+1)$ é obtida da seguinte forma:

- gera-se $\theta_1^{(j+1)}$ de $p_1(\theta_1|\theta_2^{(j)}, ..., \theta_p^{(j)})$;
- gera-se $\theta_2^{(j+1)}$ de $p_2(\theta_2|\theta_1^{(j+1)}, \theta_3^{(j)}, ..., \theta_p^{(j)})$;
- repete-se sucessivamente os passos anteriores para $i = 3, ..., p$, onde no último passo gera-se $\theta_p^{(j+1)}$ de

$$p_p(\theta_p|\theta_1^{(j+1)}, \theta_2^{(j+1)}, ..., \theta_{p-1}^{(j+1)})$$

obtendo-se o vetor $\boldsymbol{\theta}^{(j+1)} = (\theta_1^{(j+1)}, ..., \theta_p^{(j+1)})$.

Sob certas condições de regularidade, a distribuição limite de $\boldsymbol{\theta}^{(j)}$ tende a $\pi(\boldsymbol{\theta})$.

(ii) Metropolis-Hastings

O *algoritmo de Metropolis-Hastings* é utilizado para construção de cadeias de Markov e baseia-se na utilização de distribuições auxiliares.

Considera-se $q(\boldsymbol{\theta}, \cdot)$ como um núcleo arbitrário de transição e assume-se que na iteração (j) a cadeia se encontra no estado $\boldsymbol{\theta}^{(j)}$. Assim, a

posição da cadeia na iteração $(j+1)$ será denotada por $\boldsymbol{\theta}^{(j+1)}$ e é obtida a partir dos seguintes passos:

- propõe-se um movimento da cadeia para o estado $\boldsymbol{\theta}^*$ a partir de $q(\boldsymbol{\theta}^{(j)}, \cdot)$;
- aceita-se o movimento proposto com probabilidade

$$\alpha(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*) = min\left\{1, \frac{p(\boldsymbol{\theta}^*)/q(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*)}{p(\boldsymbol{\theta}^{(j)})/q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^{(j)})}\right\}$$

  e faz-se $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^*$ ou rejeita-se o movimento com probabilidade $1 - \alpha(\boldsymbol{\theta}^{(j)}, \boldsymbol{\theta}^*)$ e faz-se $\boldsymbol{\theta}^{(j+1)} = \boldsymbol{\theta}^{(j)}$, caso contrário.

Apesar do movimento da cadeia se fazer em blocos para todos os parâmetros do modelo, na prática, quando se trabalha com modelos de alta dimensão, é difícil encontrar funções $q(\boldsymbol{\theta}, \cdot)$ apropriadas. Uma variação do algoritmo comumente usada é mover uma componente do vetor de parâmetros $\boldsymbol{\theta}$ por vez. Forma-se, então, uma transição completa após o ciclo de p transições a partir de funções $q(\theta_i, \cdot)$, para $i = 1, ..., p$.

Estando decidido o método a ser utilizado, e obtida uma simulação da cadeia, deve-se verificar se a convergência foi obtida, para assim poder formar a amostra da distribuição a posteriori das quantidades desconhecidas do modelo.

Existem várias formas de se realizar uma análise a respeito da convergência da cadeia. Uma das abordagens mais comuns é a inspeção gráfica, onde analisa-se a trajetória de uma ou mais cadeias em períodos distintos de tempo. Neste estudo será utilizado este critério. O comportamento de duas ou mais cadeias com valores iniciais distintos será observado e concluir-se-á

que a convergência foi alcançada quando todas as cadeias monitoradas permanecerem em torno de um mesmo ponto. Outros critérios, mais formais, também podem ser utilizados, como os métodos propostos por Gelman e Rubin (1992) e Geweke (1992).

Após a obtenção da amostra, deve-se analisar a autocorrelação existente entre $\boldsymbol{\theta}^{(j)}$ e $\boldsymbol{\theta}^{(j+1)}$. Como estamos lidando com uma amostra de uma cadeia de Markov, temos uma amostra aleatória, mas não independente. Isto não afeta as estimativas dos parâmetros, mas tem influência sobre as variâncias das estimativas resultantes desse procedimento de amostragem (Gamerman, 1997). Assim, nos casos em que for constatada uma forte correlação serial na cadeia, após verificada a convergência, recomenda-se a retirada de uma amostra sistemática de seus valores para compor uma nova amostra. A forma como a amostragem sistemática será realizada pode ser baseada em um gráfico contendo a função de autocorrelação da cadeia.

Após a verificação destes itens, a inferência segue a partir do método de Monte Carlo. A idéia básica deste método consiste em estimar o valor de uma integral a partir da obtenção de seu valor esperado com relação a alguma distribuição de probabilidade.

3.4.2 WinBUGS

O pacote estatístico `WinBUGS` [1] é uma versão em ambiente `Windows` do pacote `BUGS` (*Bayesian Using Gibbs Sampling*). É utilizado para análise bayesiana de modelos estatísticos complexos, estimando seus parâmetros via MCMC. Foi implementado por Thomas *et al.* (1992).

[1] Disponível no site http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml

O amostrador de Gibbs implementado no `WinBugs` utiliza duas rotinas diferentes. A ARS (*Adaptive Rejection Sampling*) é usada para amostrar eficientemente qualquer distribuição condicional com função densidade log-côncava. A ARMS (*Adaptive Rejection Metropolis Sampling*) generaliza a rotina ARS para o caso de funções que não são log-côncavas.

Para o algoritmo Metropolis-Hastings, o pacote usa como densidade de transição $q(\theta^{(j)}, \cdot)$ uma distribuição gaussiana centrada no valor atual do parâmetro $\theta^{(j)}$.

Todo o processo inferencial utilizado neste trabalho foi implementado no *software* `WinBUGS versão 1.4.1`.

Capítulo 4

Modelos Lineares

Existem várias formas de se analisar a série de mortalidade por uma certa doença. Se o desejo é verificar o efeito de covariáveis na série, pode-se utilizar modelos lineares para estimar os parâmetros relacionados a cada efeito e realizar previsões.

A classe dos modelos lineares analisa a influência de covariáveis em uma determinada variável $\mathbf{Y}$ através de uma relação linear. O interesse é estudar $\mathbf{Y}$ de forma condicional aos valores de $\mathbf{X} = (\mathbf{X_1}, ..., \mathbf{X_p})$, ou seja,

$$E(\mathbf{Y}|\mathbf{X}) = \beta_1 \mathbf{X_1} + ... + \beta_p \mathbf{X_p}. \tag{4.1}$$

O modelo descrito acima é conhecido como um *modelo linear* ou *modelo linear de regressão* (Migon e Gamerman, 1999). Se é desejado a adição de um intercepto, especifica-se $\mathbf{X}_1 = 1$. Chama-se *regressão linear simples* o modelo que contém apenas uma covariável, isto é, $p = 1$.

Supondo uma amostra $\mathbf{Y} = Y_1, ..., Y_n$ independente e normalmente dis-

tribuída, a função de verossimilhança da equação (4.1) é dada por:

$$p(Y_1, ..., Y_n|\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-n} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} e_i^2\right\}$$

onde $e_i = Y_i - (\beta_1 x_{1i} + ... + \beta_p x_{pi})$, $i = 1, ..., n$.

Para estimar os valores de $\boldsymbol{\beta}$ por máxima verossimilhança, o termo $\sum_{i=1}^{n} e_i^2$ deve ser minimizado . Usando a notação matricial, temos que:

$$p(\mathbf{Y}|\boldsymbol{\beta}, \sigma^2) \propto \sigma^{-n} \exp\left\{-\frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})}{2\sigma^2}\right\}$$

onde $\mathbf{Y}' = (Y_1, ..., Y_n)$, $\mathbf{X}' = (\mathbf{x}_1', ..., \mathbf{x}_n')$ e $\boldsymbol{\beta}' = (\beta_1, .., \beta_p)$. Assim, o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ é dado por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}.$$

Na abordagem bayesiana devem ser especificadas distribuições a priori para os parâmetros do modelo.

Em geral, são especificadas distribuições a priori Normal($\mathbf{A}$, $\phi^{-1}\mathbf{B}$) para o vetor $\boldsymbol{\beta}$ e distribuições Gama(c, d) para o parâmetro $\phi = \sigma^{-2}$. Obtém-se, então, distribuições a posteriori t-Student e Gama para $\boldsymbol{\beta}|\mathbf{Y}$ e $\phi|\mathbf{Y}$, respectivamente.

4.1 Modelos Lineares Generalizados

Modelar o número de mortes por DISC em um determinado período de tempo por meio de um modelo linear nem sempre é recomendado, uma vez que estes modelos supõem normalidade da série temporal. Como a série de números de mortos pela doença é discreta, uma outra forma de relacionamento linear entre as covariáveis e a série de dados é utilizada.

Uma extensão dos modelos lineares permite modelar observações descritas por membros da família exponencial. Esta classe de modelos é conhecida como *Modelos Lineares Generalizados* (MLG). Uma descrição detalhada destes modelos pode ser encontrada em McCullagh e Nelder (1989).

De uma forma geral, a estrutura de um MLG é formada por três partes: uma *componente aleatória* composta de uma variável aleatória $\mathbf{Y}$ com n observações independentes, um vetor de médias μ e uma distribuição pertencente à família exponencial; uma *componente sistemática* composta por variáveis explicativas $(x_1, ..., x_p)$ tais que produzem um preditor linear η; e uma função monótona diferenciável, conhecida como *função de ligação*, que relaciona estas duas componentes (Cordeiro e Lima Neto, 2004). Esta estrutura é detalhada a seguir.

Seja $\mathbf{Y}' = (Y_1, ..., Y_n)$ independentes e identicamente distribuídas, com médias $\boldsymbol{\mu}' = (\mu_1, ..., \mu_n)$. Suponha que cada componente de $\mathbf{Y}$ possui distribuição pertencente à família exponencial, na forma a seguir:

$$f_Y(y; \theta, \phi) = \exp\left\{\frac{[y\theta - b(\theta)]}{a(\phi)} + c(y, \phi)\right\} \tag{4.2}$$

onde $a(\cdot)$, $b(\cdot)$ e $c(\cdot)$ são funções conhecidas; $\phi > 0$ é o parâmetro de dispersão e θ é o parâmetro canônico da distribuição.

Sabe-se que a função $b(\theta)$ é duas vezes diferenciável em θ e que

$$\mu = E[Y] = b'(\theta).$$

Assim, μ e o parâmetro θ se relacionam através da transformação inversa, supondo que esta exista,

$$\theta = b'^{-1}(\mu).$$

No trabalho em estudo, uma escolha natural é modelar o número de

mortes anuais por faixa etária x, D_{xt}, de acordo com uma distribuição de Poisson.

Neste caso, tem-se que:

$$p(D_{xt}|\lambda_{xt}) = \frac{e^{-\lambda_{xt}}(\lambda_{xt})^{d_{xt}}}{d_{xt}!} = \exp\left\{[d_{xt}\log(\lambda_{xt}) - \lambda_{xt}] - \log(d_{xt}!)\right\}, d_{xt} = 0, 1, 2, ...$$

onde $\theta_{xt} = \log(\lambda_{xt})$, $a(\phi) = 1$, $b(\theta_{xt}) = \lambda_{xt}$ e $c(d_{xt}, \phi) = -\log(d_{xt}!)$.

A função de ligação entre a média da série λ_{xt} e o parâmetro θ_{xt} é a função $\exp(\cdot)$. O preditor linear seria, então, da seguinte forma:

$$\theta_{xt} = \log(\lambda_{xt}) = \mathbf{X}'_{\mathbf{xt}}\boldsymbol{\beta}.$$

Na inferência clássica, os parâmetros $\beta's$ podem ser estimados por máxima verossimilhança, aplicando-se testes para sua significância e predizendo os valores da série. Na abordagem bayesiana, são especificadas distribuições a priori para o vetor de parâmetros $\boldsymbol{\beta}$ e faz-se a inferência através da análise da distribuição a posteriori $p(\boldsymbol{\beta}|y_{xt})$, obtida através da atualização da informação via Teorema de Bayes.

4.2 Modelos Lineares Dinâmicos

A classe dos *Modelos Lineares Dinâmicos* (MLD) é muito utilizada para a análise de séries temporais e regressão. Estes modelos atualizam os parâmetros de forma sequencial com base nas observações passadas. Uma análise desta classe é detalhada em West e Harrison (1997).

Os modelos lineares dinâmicos são modelos onde a variação do parâmetro no tempo e os dados disponíveis são descritos probabilisticamente (Migon e

Gamerman, 1999). Estes modelos são caracterizados pela *equação observacional* (4.3) e *pela equação de sistema* (4.4), que apresentam a seguinte forma:

$$Y_t = \mathbf{x}_t' \boldsymbol{\beta}_t + \epsilon_t, \quad \epsilon_t \sim Normal(0, \sigma_t^2) \tag{4.3}$$

$$\boldsymbol{\beta}_t = \mathbf{G_t} \boldsymbol{\beta}_{t-1} + \mathbf{w}_t, \quad \mathbf{w}_t \sim Normal(\mathbf{0}, \mathbf{W}_t) \tag{4.4}$$

onde Y_t é a série, $\mathbf{x}_t'$ um vetor de variáveis explicativas, $\boldsymbol{\beta}_t$ os parâmetros do modelo, $\mathbf{G_t}$ uma matriz representando a evolução dos parâmetros no tempo, σ_t^2 a variância do erro associado à observação univariada e $\mathbf{W}_t$ a variância dos erros associados ao vetor de parâmetros.

Esta classe de modelos fica especificada por meio da quádrupla

$$\left\{ \mathbf{x}_t, \mathbf{G_t}, \sigma_t^2, \mathbf{W}_t \right\}.$$

Nota-se que a distribuição de Y_t é independente das observações passadas dado o conhecimento de $\boldsymbol{\beta}_t$. Isto significa que a dinâmica temporal é sumarizada no processo de evolução dos parâmetros.

Na abordagem bayesiana, realiza-se predições e inferências para os parâmetros através da utilização da informação atualizada no tempo $t-1$. Assim, a distribuição a posteriori em $t-1$ é utilizada pelo modelo como distribuição a priori em t.

A utilização da classe dos Modelos Lineares Dinâmicos em séries temporais permite que mudanças bruscas sejam suavizadas além de permitir a acomodação natural de informações subjetivas (Ferreira e Gamerman, 1998). Entretanto, a utilização destes modelos no presente estudo não é recomendada pois os dados são discretos.

4.3 Modelos Lineares Generalizados Dinâmicos

West *et al.* (1985) extenderam a classe dos Modelos Lineares Dinâmicos para dados que pertençam à família exponencial.

A classe dos *Modelos Lineares Generalizados Dinâmicos* (MLGD) é caracterizada pelas seguintes componentes:

- Modelo observacional: a equação observacional do MLD é substituída pelo modelo observacional conforme apresentado em (4.2);
- Função de Ligação: relaciona a média da série com os parâmetros do modelo

  $$g(\boldsymbol{\mu}_t) = \boldsymbol{\theta}_t = \mathbf{x}_t'\boldsymbol{\beta}_t,$$

  onde

  $$\boldsymbol{\mu}_t = E[\mathbf{Y}_t|\boldsymbol{\theta}_t] = b'(\boldsymbol{\theta}_t).$$

- Equação de evolução:

  $$\boldsymbol{\beta}_t = \mathbf{G_t}\boldsymbol{\beta}_{t-1} + \mathbf{w}_t, \mathbf{w}_t \sim Normal(\mathbf{0}, \mathbf{W}_t)$$

O processo de inferência é equivalente ao MLD. Porém, este processo é prejudicado pois o cálculo das integrais necessárias para a atualização dos parâmetros não pode ser obtido exatamente. Vários métodos para lidar com este problema foram propostos como, por exemplo, os métodos de MCMC.

4.4 Modelos Lineares Hierárquicos

Assim como é possível especificar distribuições a priori para os parâmetros de um modelo em estágios, pode-se também utilizar esta abordagem no processo de modelagem de fenômenos.

A título de ilustração de uma modelagem em dois estágios, suponha que $\mathbf{Y}$ é uma variável de interesse, e que deseja-se analisá-la através de um modelo linear. Desta forma, temos que

$$\mathbf{Y}|\boldsymbol{\beta}, \phi \sim Normal(\mathbf{X}\boldsymbol{\beta}, \phi^{-1}I),$$

onde $\boldsymbol{\beta} = (\beta_1, ..., \beta_p)$ é o vetor de parâmetros e $\mathbf{X}$ é uma matriz de covariáveis.

Supondo que exista uma matriz $\mathbf{X}^*$ que contenha covariáveis que explicam o vetor de parâmetros $\boldsymbol{\beta}$, pode-se incluir o seguinte estágio no modelo:

$$\boldsymbol{\beta}|\boldsymbol{\beta}^*, \phi \sim Normal(\mathbf{X}^*\boldsymbol{\beta}^*, \phi^{-1}C_1),$$

onde $\boldsymbol{\beta}^*$ é um vetor de parâmetros associado às covariáveis da matriz $\mathbf{X}^*$. Completando o modelo,

$$\boldsymbol{\beta}^*|\phi \sim Normal(\boldsymbol{\mu}, \phi^{-1}C_2)$$

$$\phi \sim Gama(a, b).$$

Assim, preserva-se a linearidade do modelo e a inferência é realizada de forma simples, uma vez que esta modelagem produz distribuições a posteriori conhecidas. Dependendo do problema a ser analisado, mais estágios podem ser adicionados no modelo. Porém, quando utilizados muitos estágios, torna-se mais difícil a interpretação dos parâmetros envolvidos.

Capítulo 5

Modelos Propostos

Neste capítulo são apresentadas as características gerais dos diferentes modelos propostos para a análise dos dados deste trabalho, os quais baseiam-se nas idéias de Modelos Generalizados Dinâmicos (West *et al.*, 1985), Modelos Hierárquicos e no conceito de Fatores de Redução de mortalidade.

5.1 Modelagem do número de óbitos por DISC

Para descrição da relação entre número observado de mortes ao longo dos anos por doenças isquêmicas do coração e as classes etárias correspondentes, assumindo que todos os indivíduos pertecentes ao mesmo sexo e a mesma classe etária morrem por DISC independentemente e com a mesma probabilidade, será utilizado o seguinte modelo linear generalizado:

$$D_{xt}|\mu_{xt} \sim Poisson(\lambda_{xt}) \tag{5.1}$$

$$E(D_{xt}) = \lambda_{xt} = e_{xt}.\mu_{xt}, \quad x = x_{inf}, \ldots, x_{sup}, t = 1, \ldots, T \tag{5.2}$$

onde D_{xt} é o número observado de mortos por doenças isquêmicas do coração para a classe etária x e o tempo t; μ_{xt} é a força de mortalidade para a classe etária x e o tempo t, $\mu_{xt} > 0$; e_{xt} é o número de expostos ao risco para a classe etária x e o tempo t, o qual é supostamente conhecido.

A partir da função de ligação canônica (η_{xt}) da distribuição Poisson, temos que:

$$\eta_{xt} = \log(\lambda_{xt}) = \log(e_{xt}.\mu_{xt}) = \log(e_{xt}) + \log(\mu_{xt}) \tag{5.3}$$

Assim, nosso interesse será modelar $\log(\mu_{xt})$, isto é, a força de mortalidade para a classe etária x no tempo t na escala logarítmica utilizando uma abordagem bayesiana a fim de incorporar a incerteza inicial associada aos parâmetros do modelo no processo de estimação.

5.2 Modelo Dinâmico Local

Neves e Migon (2007) modelaram a evolução das forças de mortalidade usando um modelo dinâmico generalizado (West *et al.*, 1985), considerando que as diferentes forças de mortalidade evoluem no tempo e nas faixas etárias por meio de perturbações multiplicativas:

$$\log(\mu_{xt}) = \log(\mu_{xt-1}) + \omega_t, \tag{5.4}$$

$$\omega_t \sim Normal(0, W_t), \quad x = x_{inf}, \ldots, x_{sup}, t = 2, \ldots, T$$

onde W_t é modelado a priori como uma distribuição Gama Inversa não informativa. Especificou-se distribuições a priori Gama não informativas para os parâmetros μ_{x1}. Para os demais anos, tem-se então que:

$$\mu_{xt} \sim Gama(a, b) I_{(\mu_{x-1t}, \mu_{x+1t})}(\mu_{xt}), x = x_{inf}, \ldots, x_{sup} \tag{5.5}$$

onde $t = 1$, $I_A(y) = 1$, se $y \in A$ e $I_A(y) = 0$, caso contrário. Adicionalmente, assume-se que os hiperparâmetros a e b sejam conhecidos.

5.3 Modelo Dinâmico Global

Estendendo a Lei de Gompertz para que seja incluída a dimensão temporal, na escala logarítmica (2.7), temos que:

$$\log(\mu_{xt}) = \alpha_t + \beta_t.x, \quad x = x_{inf}, \ldots, x_{sup}, t = 1, \ldots, T. \tag{5.6}$$

A partir daí, reescrevemos a função de ligação canônica dada em (5.3) para:

$$\eta_{xt} = \log(e_{xt}) + \log(\mu_{xt}) = \log(e_{xt}) + \alpha_t + \beta_t.x, \quad x = x_{inf}, \ldots, x_{sup}, t = 1, \ldots, T \tag{5.7}$$

onde $\mu_{xt} > 0$ e e_{xt} são constantes conhecidas.

Assim, a proposta é modelar o $\log(\mu_{xt})$ por meio da equação (5.6), tendo como covariável a classe etária em anos. A evolução temporal dos parâmetros é descrita a seguir:

$$\alpha_t = \alpha_{t-1} + \omega a_t \tag{5.8}$$

$$\beta_t = \beta_{t-1} + \omega b_t, \quad t = 2, \ldots, T \tag{5.9}$$

onde $\omega z_t \sim Normal(0, Wz_t)$, $z = a, b$ e Wz_t são modeladas por meio de distribuições Gamas Inversas não informativas. Finalmente, especifica-se distribuições a priori Normais não informativas para os parâmetros α_1 e β_1, satisfazendo as restrições impostas pela Lei de Gompertz.

5.4 Modelo Hierárquico

Ao estudar as doenças isquêmicas do coração no estado do Rio de Janeiro no mesmo período considerado, Tura (2006) observou comportamento linear dos parâmetros $B = \exp(\alpha)$ e $log(c) = \beta$ no tempo e, com base na equação (5.6), propôs o seguinte modelo hierárquico:

$$\begin{aligned} \log(\mu_{xt}) &= \alpha_t + \beta_t.x \\ \alpha_t &= a_0 + a_1.t + \varepsilon_1 \\ \beta_t &= b_0 + b_1.t + \varepsilon_2 \end{aligned} \tag{5.10}$$

onde $\varepsilon_i \sim Normal(0, \sigma_i^2)$, $i = 1, 2$.

Partindo do mesmo princípio, especificamos quatro diferentes estruturas para α_t e β_t, $t = 1, \ldots, T$, a partir da especificação de quatro diferentes prioris, a saber:

Priori I:

$\alpha_t \mid a_0, \sigma_1^2 \sim Normal(a_0, \sigma_1^2), \quad t = 1, \ldots, T$

$\beta_t \mid b_0, \sigma_2^2 \sim Normal(b_0, \sigma_2^2), \quad t = 1, \ldots, T$

Priori II:

$\alpha_t \mid a_0, a_1, \sigma_1^2 \sim Normal(a_0 + a_1 t, \sigma_1^2), \quad t = 1, \ldots, T$

$\beta_t \mid b_0, \sigma_2^2 \sim Normal(b_0, \sigma_2^2), \quad t = 1, \ldots, T$

Priori III:

$\alpha_t \mid a_0, \sigma_1^2 \sim Normal(a_0, \sigma_1^2), \quad t = 1, \ldots, T$

$\beta_t \mid b_0, b_1, \sigma_2^2 \sim Normal(b_0 + b_1 t, \sigma_2^2), \quad t = 1, \ldots, T$

Priori IV:

$\alpha_t \mid a_0, a_1, \sigma_1^2 \sim Normal(a_0 + a_1 t, \sigma_1^2), \quad t = 1, \ldots, T$

$\beta_t \mid b_0, b_1, \sigma_2^2 \sim Normal(b_0 + b_1 t, \sigma_2^2), \quad t = 1, \ldots, T$

considerando que $\tau_i = 1/\sigma_i^2$ possui distribuição a priori Gama não-informativa e os hiperparâmetros a_i e b_i, $i = 1, 2$ tendo distribuição Normal também não-informativa.

5.5 Modelo Linear Generalizado com Fator de Redução de Mortalidade

Renshaw e Haberman (2003) propuseram modelar o número de mortes D_{xt} como variáveis respostas independentes de um modelo linear generalizado Poisson com sobredispersão, tendo o tempo como covariável conhecida, e utilizando o conceito de Fator de Redução de Mortalidade, desenvolvido por atuários britânicos.

5.5.1 Fator de Redução de Mortalidade

O fator de redução de mortalidade F é definido em função da força de mortalidade μ_{xt}, em uma classe etária x e no tempo t, e também em função de uma tábua de mortalidade padrão válida em um instante t_0, sendo caracterizada como:

$$\mu_{xt} = \mu_{xt_0}.F(x, t), \qquad t \geq t_0, \tag{5.11}$$

sujeita à restrição:

$$F(x, t_0) = 1, \qquad \forall x \tag{5.12}$$

onde μ_{xt_0} a força de mortalidade para idade x, em um ano padrão t_0. Adicionalmente, define-se também que F é monótona decrescente com o aumento de t, $\forall x$ e $t > t_0$, assim

$$0 < F(x, t) \leq 1, \qquad \forall x, t > t_0. \tag{5.13}$$

5.5.2 O Modelo

Um modelo para a força de mortalidade é o que a coloca como função do fator de redução de mortalidade, como descrito em (5.11).

Assim, reescrevendo (5.2) temos que:

$$E(D_{xt}) = \lambda_{xt} = e_{xt}.\mu_{xt} = e_{xt}.\mu_{xt_0}.F(x, t) \tag{5.14}$$

$$var(D_{xt}) = \phi E(D_{xt}), \tag{5.15}$$

onde ϕ é o parâmetro de sobredispersão.

Desta forma, o preditor linear é:

$$\eta_{xt} = log(e_{xt}) + log(\mu_{xt_0}) + log\{F(x, t)\}, \tag{5.16}$$

sujeito à restrição $log\{F(x, t_0)\} = 0$ e tendo como *offset* $log(e_{xt}) + log(\mu_{xt_0})$.

Os fatores de redução de mortalidade podem ser modelados segundo uma das seguintes equações:

$$log\{F(x, t)\} = \gamma_t(t - t_0) + \beta_x(t - t_0) \tag{5.17}$$

$$log\{F(x, t)\} = \gamma_t(t - t_0) + \beta'_x(t - t_j)_- + \beta_x(t - t_0) \tag{5.18}$$

para $t \in [t_1, t_n]$, com $t_1 < t_j < t_0 \leq t_n$, $(t)_- = t$, $t \leq 0$ e $(t)_- = 0$, $t > 0$.

Na equação (5.18) a adição do ponto de quebra t_j, onde $t_j < t_0$, introduz maior flexibilidade estrutural. Os termos γ_t são adicionados para garantir a igualdade entre o número de mortes observado e esperado para cada tempo t, e a motivação é o fato do modelo linear generalizado Poisson ser próximo da análise de tabelas de contingência. Estes termos não influenciam na projeção.

Uma vez que o modelo será analisado sob o enfoque bayesiano, existe a necessidade de especificar-se distribuições a priori para os parâmetros do modelo. Aos parâmetros β_x, β'_x e $\log(\phi)$ são alocadas distribuições a priori Normais não informativas, respeitando suas respectivas restrições. Os parâmetros γ_t são modelados por meio de duas diferentes prioris Normais: não informativa e com uma estrutura dinâmica na média, isto é, dependendo do γ no tempo anterior.

Capítulo 6

Resultados

Neste capítulo são apresentados os principais resultados obtidos da aplicação dos modelos descritos no capítulo anterior aos dados de mortalidade por DISC no estado do Rio de Janeiro no período de 1980 a 2002.

6.1 Descrição dos Dados

Este trabalho foi realizado com os dados referentes as séries históricas da população do estado do Rio de Janeiro e o número de óbitos por doenças isquêmicas do coração neste mesmo estado, entre os anos de 1980 e 2002, separadas por gênero, de adultos de mais de 20 anos.

Os dados referentes à população residente no estado do Rio de Janeiro foram obtidos em consulta ao Instituto Brasileiro de Geografia e Estatística (IBGE). Estes são referentes a população em 1º de julho de cada ano e são divididos nas seguintes classes etárias: 20 a 24 anos, 25 a 29 anos, 30 a 34

anos, 35 a 39 anos, 40 a 44 anos, 45 a 49 anos, 50 a 54 anos, 55 a 59 anos, 60 a 64 anos, 65 a 69 anos, 70 a 74 anos, 75 a 79 anos e 80 anos ou mais.

Os dados referentes à mortalidade por DISC foram obtidos em consulta ao Sistema de Informações sobre Mortalidade (SIM), do DATASUS/MS, utilizando-se a 9ª Conferência de Revisão da Classificação Internacional de Doenças de 1975 (CID 9), para óbitos de 1980 a 1995, com os códigos entre 410 e 414. Para os óbitos a partir de 1996 foi utilizada a 10ª Conferência de Revisão da Classificação Internacional de Doenças de 1995 (CID 10) com os códigos entre I20 e I25.

Os Gráficos 6.1 e 6.2 apresentam a evolução das taxas de mortalidade observadas por DISC, na escala logarítmica, por classe etária, para homens e mulheres, respectivamente.

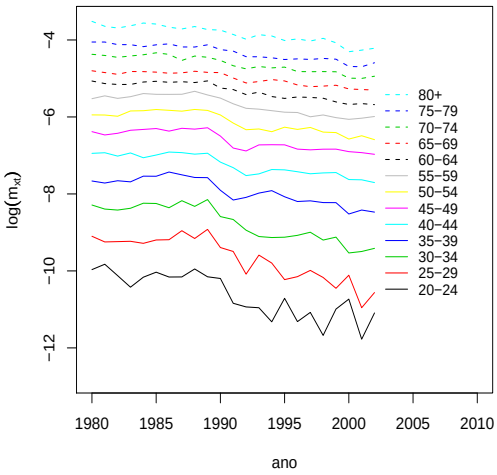

Figura 6.1: Taxa bruta de mortalidade por DISC na escala logarítmica por ano para cada classe etária para o sexo masculino.

Analisando os gráficos 6.1 e 6.2, percebe-se que:

- a mortalidade é mais acentuada para idades mais avançadas, conforme já citado na Introdução;
- as taxas de mortalidade estão decrescendo para todas as faixas etárias no período considerado;
- ao final dos anos 80 o efeito de redução dessas taxas é intensificado sugerindo possível mudança estrutural neste declínio.

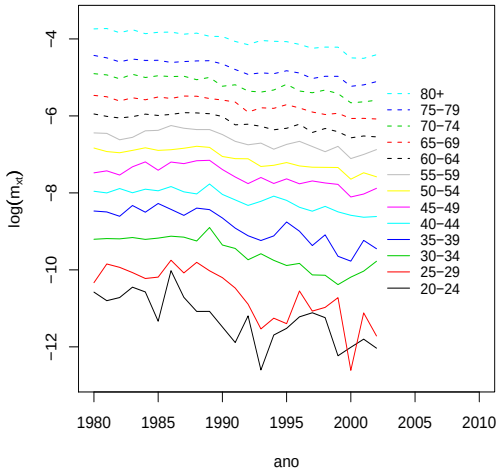

Figura 6.2: Taxa bruta de mortalidade por DISC na escala logarítmica por ano para cada classe etária para o sexo feminino.

Analisando as taxas por gênero, nota-se que a mortalidade por DISC entre as mulheres é menor que a dos homens praticamente em todo o período considerado e para todas as faixas etárias.

Os gráficos 6.3 e 6.4 apresentam a evolução das taxas de mortalidade por DISC, na escala logarítmica, para os anos 1980, 1991 e 2000, para homens e mulheres, respectivamente.

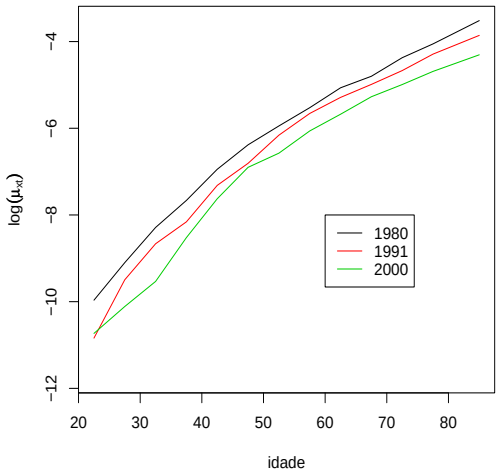

Figura 6.3: Taxa bruta de mortalidade por DISC na escala logarítmica por idade para os anos 1980, 1991 e 2000, para o sexo masculino.

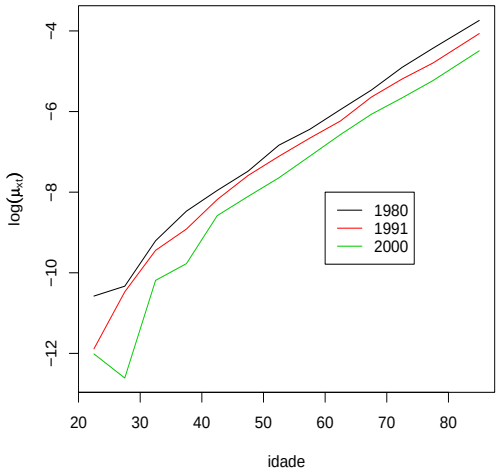

Figura 6.4: Taxa bruta de mortalidade por DISC na escala logarítmica por idade para os anos 1980, 1991 e 2000, para o sexo feminino.

Analisando as figuras 6.3 e 6.4 nota-se redução das taxas de mortalidade ao comparar-se os anos 1980, 1991 e 2000. Esta redução das taxas é evidente para ambos os sexos, sendo que nas mulheres, percebe-se uma mudança mais significativa nas faixas etárias mais baixas.

6.2 Modelo Dinâmico Local

O modelo dinâmico local, conforme apresentado na seção 5.2, foi ajustado aos dados utilizando-se as seguintes prioris:

- $W_t = \frac{1}{\tau_t}$ com $\tau_t \sim$ Gama(10,10), $\forall t$;
- para $x = 1$: $\mu_{1,1} \sim$ Gama(0.001,0.001) I(0,$\mu_{2,1}$);
- para $x = 2, \ldots, 12$: $\mu_{x,1} \sim$ Gama(0.001,0.001) I($\mu_{x-1,1}$,$\mu_{x+1,1}$);
- para $x = 13$: $\mu_{13,1} \sim$ Gama(0.001,0.001) I($\mu_{12,1}$,1).

A inferência para o modelo dinâmico local foi realizada utilizando-se o algoritmo MCMC, obtendo-se amostras das distribuições a posteriori dos parâmetros. Este método foi implementado no *software* `WinBUGS`.

A convergência das cadeias foi monitorada graficamente por meio da análise das trajetórias de duas cadeias partindo de valores iniciais distintos. A fim de reduzir a autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori apenas 1 (um) valor a cada 10 iterações. A partir disto, concluiu-se que a convergência havia sido obtida após amostrarmos 500 valores para o modelo considerado.

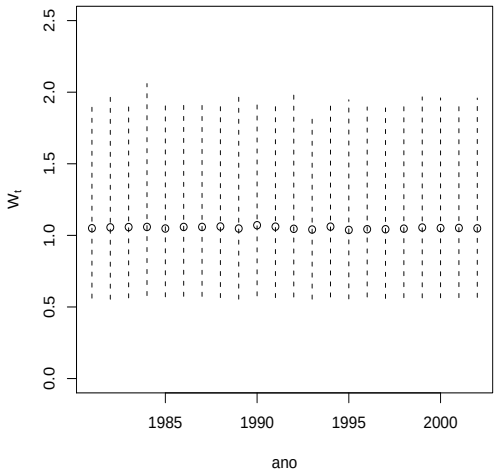

Figura 6.5: Médias a posteriori e respectivos intervalos de 95% de credibilidade de W_t para o modelo dinâmico local para o sexo masculino.

Para compor a amostra, observou-se as cadeias por mais 1000 iterações, totalizando duas amostras de tamanho 1000 para cada parâmetro.

As figuras 6.5 e 6.6 apresentam as médias a posteriori e respectivos intervalos de 95% de credibilidade de W_t para o modelo dinâmico local para os sexo masculino e feminino, respectivamente.

Observando o gráfico 6.5 notamos que as estimativas para o parâmetro W_t e seus respectivos intervalos de credibilidade sugerem que não existem diferenças significativas entre as variâncias dos efeitos aleatórios do modelo dinâmico local no tempo, para o sexo masculino. Em outras palavras, o passeio aleatório que descreve as mudanças no nível das séries das forças de mortalidade na escala logarítmica para os homens parece possuir variância em torno de 1.05.

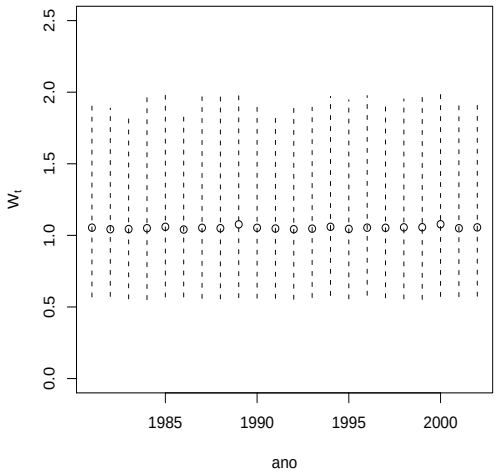

Figura 6.6: Médias a posteriori e respectivos intervalos de 95% de credibilidade de W_t para o modelo dinâmico local para o sexo feminino.

Conclusão semelhante é obtida analisando-se os resultados para as mulheres, conforme apresentado na figura 6.6.

Tendo em vista os resultados obtidos para ambos os sexos, considerou-se apropriado ajustar um modelo com W_t constante. Assim, realizou-se novamente a inferência utilizando esta simplificação do modelo. Nesta implementação, foram novamente monitoradas duas cadeias partindo-se de valores iniciais distintos, sendo que foram considerados na amostra apenas 1 a cada 20 valores, devido à autocorrelação. Ao fim de 250 iterações nestas condições, considerou-se, por meio de inspeção visual, que a convergência foi alcançada e, após isso, foram ainda amostrados mais 1000 valores para composição da amostra.

Os histogramas das distribuições a posteriori de W, para homens e mulheres, são apresentados nas figuras 6.7 e 6.8.

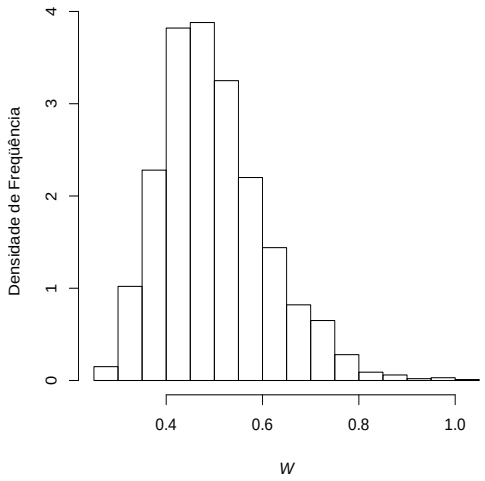

Figura 6.7: Histograma da distribuição a posteriori do parâmetro W para o modelo dinâmico local para o sexo masculino.

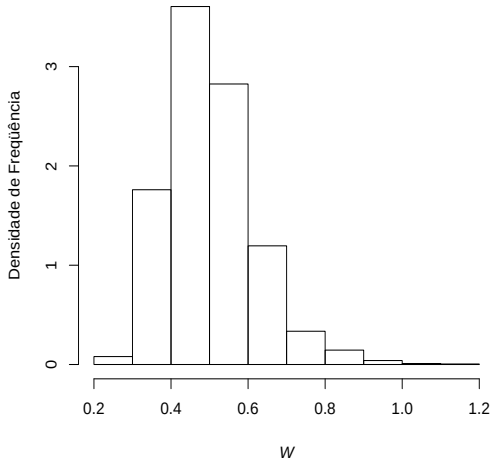

Figura 6.8: Histograma da distribuição a posteriori do parâmetro W para o modelo dinâmico local para o sexo feminino.

Os modelos com W constante e variando ao longo do tempo serão comparados para verificar qual produz o melhor ajuste. O critério de seleção adotado destacará o modelo que apresentar o menor valor do DIC - *Deviance Information Criterion* (Spiegelhalter *et al.*, 2002). O Apêndice apresenta uma breve descrição deste critério.

A tabela 6.1 apresenta os valores dos DIC's para os modelos locais considerados para os sexos masculino e feminino.

Tabela 6.1: Comparação dos modelos dinâmicos considerando W variando ao longo do tempo e constante, por meio do DIC, para ambos os sexos.

Sexo	W	$\bar{D}$	$\hat{D}$	p_D	DIC
Masculino	Varia no tempo	3128.050	3092.830	35.216	3163.270
Masculino	**Constante**	3127.580	3092.820	34.765	**3162.350**
Feminino	Varia no tempo	2811.130	2775.840	35.295	2846.430
Feminino	**Constante**	2810.760	2775.790	34.976	**2845.740**

De acordo com o critério DIC, o melhor modelo dinâmico local é o que considera o parâmetro W como sendo constante, isto é, não variando ao longo do tempo, para ambos os sexos.

As figuras 6.9 e 6.10 apresentam as estimativas, obtidas a partir das médias a posteriori, das forças de mortalidade na escala logarítmica para o modelo local com W constante, para homens e mulheres, respectivamente.

Analisando a figura 6.9 observa-se que as médias a posteriori das log-forças de mortalidade apresentam reduções mais acentuadas no início dos anos 90 e no ano 2000, sendo que este último está de acordo com o *efeito de*

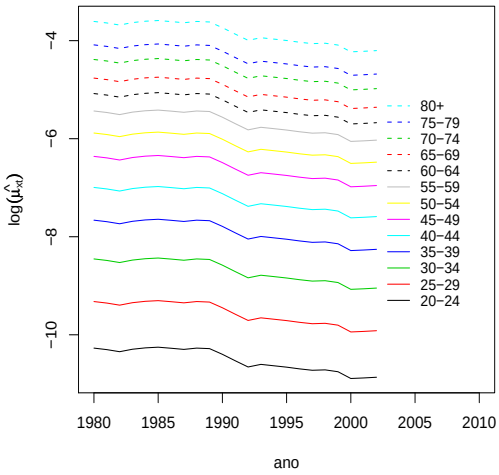

Figura 6.9: Médias a posteriori de $\log(\mu_{xt})$ para o modelo dinâmico local com W_t constante para o sexo masculino.

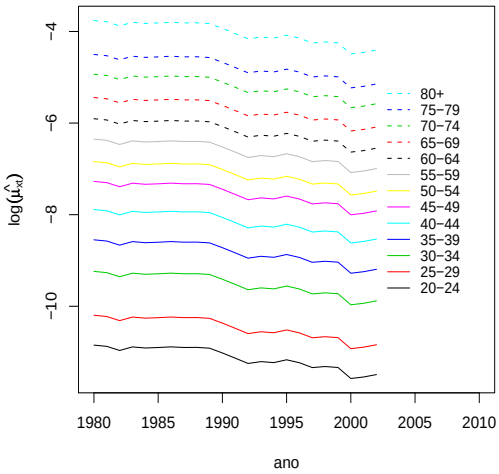

Figura 6.10: Médias a posteriori de $\log(\mu_{xt})$ para o modelo dinâmico local com W_t constante para o sexo feminino.

período apontado na seção 1.2. O mesmo efeito é observado analisando-se os resultados para o sexo feminino, na figura 6.10. Adicionalmente, analisando-se ambas as figuras, confirma-se que as mortalidades do sexo masculino encontram-se em patamares mais elevados que as do sexo feminino.

6.3 Modelo Dinâmico Global

O modelo dinâmico global, conforme apresentado na seção 5.3, foi ajustado aos dados utilizando-se as seguintes prioris:

- $Wz_t = \frac{1}{\tau z_t}$ com $\tau z_t \sim$ Gama(1,1), $z = a, b$, $\forall t$;
- $\alpha_1 \sim$ Normal(0,100);
- $\beta_1 \sim$ Normal(0,100)I(0,∞).

A inferência para o modelo dinâmico global foi realizada utilizando-se o algoritmo MCMC, obtendo-se amostras das distribuições a posteriori dos parâmetros. Este método foi implementado no *software* `WinBUGS`.

A convergência das cadeias foi monitorada graficamente por meio da análise das trajetórias de duas cadeias partindo de valores iniciais distintos. A fim de reduzir a autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori apenas 1 (um) valor a cada 10 iterações. A partir disto, concluiu-se que a convergência havia sido obtida após amostrarmos 500 valores para o modelo considerado.

Para compor a amostra, observou-se as cadeias por mais 1000 iterações, totalizando duas amostras de tamanho 1000 para cada parâmetro.

As figuras 6.11 e 6.12 apresentam as médias a posteriori e respectivos intervalos de 95% de credibilidade de Wa_t para o modelo dinâmico global para homens e mulheres, respectivamente.

Analisando-se a figura 6.11, pode-se perceber que a estimativa de Wa_t para o ano de 1992 para os homens é um pouco mais elevada do que àquelas

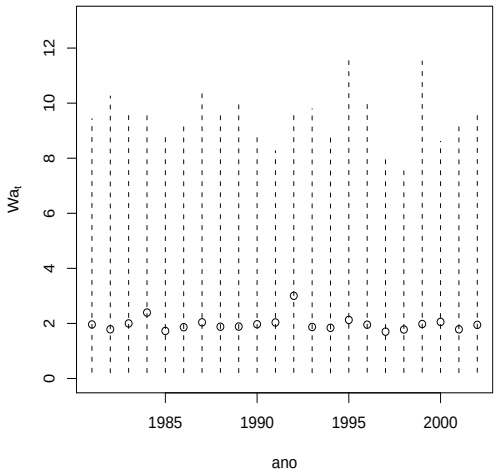

Figura 6.11: Médias a posteriori e respectivos intervalos de 95% de credibilidade de Wa_t para o modelo dinâmico global para o sexo masculino.

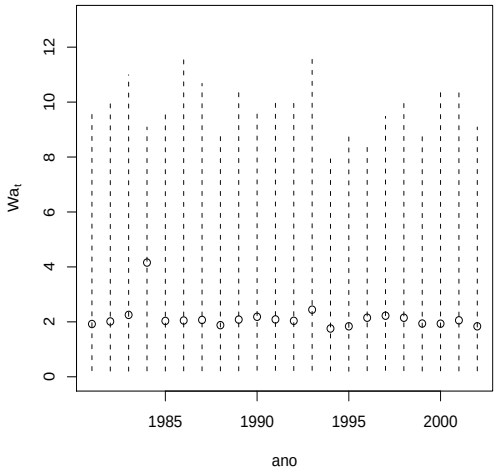

Figura 6.12: Médias a posteriori e respectivos intervalos de 95% de credibilidade de Wa_t para o modelo dinâmico global para o sexo feminino.

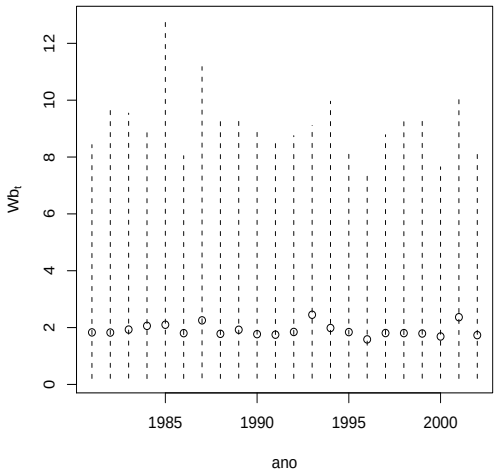

Figura 6.13: Médias a posteriori e respectivos intervalos de 95% de credibilidade de Wb_t para o modelo dinâmico global para o sexo masculino.

associadas aos demais anos. O mesmo acontece no ano de 1984 para as mulheres, sendo que neste caso o valor é ainda mais elevado em comparação aos demais anos.

A evolução das médias a posteriori de Wb_t para homens e mulheres, e respectivos intervalos de 95% de credibilidade, são apresentadas nas figuras 6.13 e 6.14. Nota-se que no modelo para as mulheres a estimativa para o ano de 1980 foi superior aos demais anos.

A evolução temporal das médias a posteriori dos parâmetros α_t, e seus respectivos intervalos de 95% de credibilidade, são apresentadas nas figuras 6.15 e 6.16, para os sexos masculino e feminino.

A trajetória das médias a posteriori de β_t, e seus respectivos intervalos de 95% de credibilidade, são apresentadas nas figuras 6.17 e 6.18, para ambos

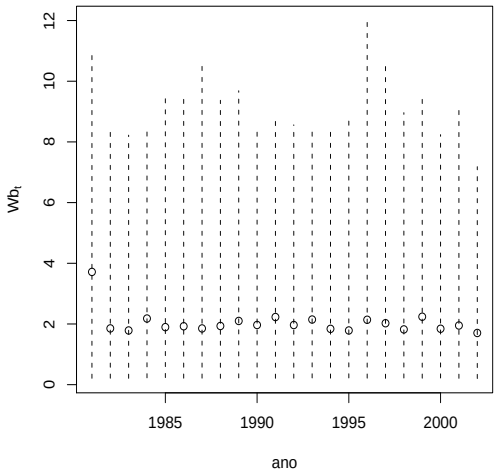

Figura 6.14: Médias a posteriori e respectivos intervalos de 95% de credibilidade de Wb_t para o modelo dinâmico global para o sexo feminino.

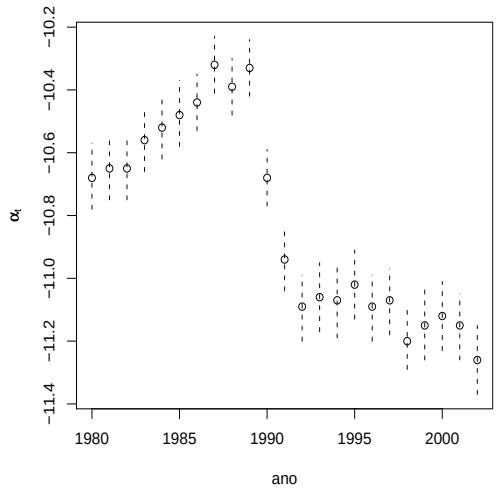

Figura 6.15: Médias a posteriori e respectivos intervalos de 95% de credibilidade de α_t para o modelo dinâmico global para o sexo masculino.

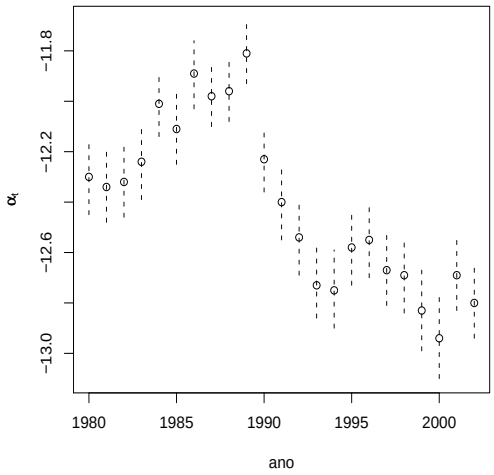

Figura 6.16: Médias a posteriori e respectivos intervalos de 95% de credibilidade de α_t para o modelo dinâmico global para o sexo feminino.

os sexos.

Analisando-se as médias a posteriori de α_t e de β_t, para ambos os sexos, observa-se que de 1980 a 1989 o aumento do nível da série (α_t) é compensado por uma redução gradual do efeito das idades nas reduções das mortalidades. Por outro lado, de 1990 a 1992, o inverso ocorre. A partir de 1992 as mudanças são mais suaves para α_t enquanto que β_t sofre uma elevação rápida retornando aos patamares do início da série. Com base nestas observações, pode-se concluir que a redução das mortalidades por DISC primeiramente ocorreu com uma redução das diferenças entre as mortalidades das faixas etárias durante a década de 80. Em seguida, no início da década de 90, as diferenças entre as mortalidades para as diferentes faixas etárias voltaram a elevar-se, mantendo-se em torno deste patamar a partir de 1992, à exceção de uma redução ocorrida no ano 2000.

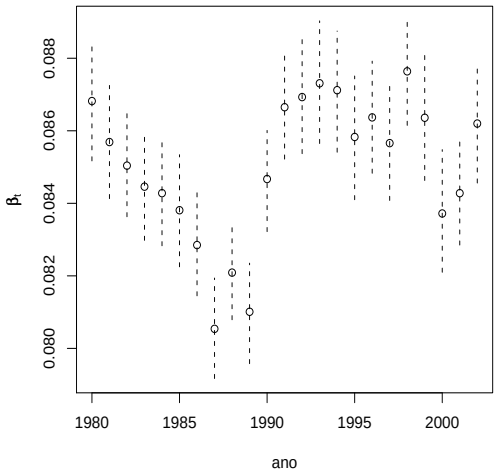

Figura 6.17: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_t para o modelo dinâmico global para o sexo masculino.

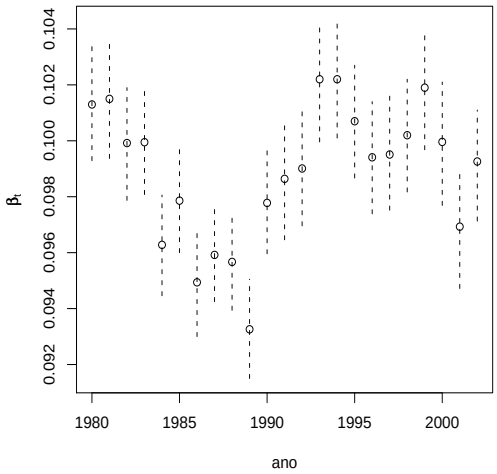

Figura 6.18: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_t para o modelo dinâmico global para o sexo feminino.

Os valores dos DIC's para o modelo dinânimo global são apresentados na tabela 6.2.

Analisando-se os valores obtidos para o DIC, observa-se que estes são muito superiores aos obtidos pelo modelo dinâmico local, em especial para o sexo masculino.

Tabela 6.2: Valores dos DIC's para o modelo dinâmico global, para ambos os sexos.

Sexo	$\bar{D}$	$\hat{D}$	p_D	DIC
Masculino	11772.900	11727.300	45.586	11818.500
Feminino	3578.320	3532.250	46.064	3624.380

6.4 Modelos Hierárquicos

Os modelos hierárquicos, conforme apresentados na seção 5.4, foram ajustados aos dados utilizando as prioris descritas na referida seção, onde

- $a_i \sim Normal(0, 1000)$, $i = 0, 1$;
- $b_i \sim Normal(0, 1000)$, $i = 0, 1$;
- $\tau_i = \frac{1}{\sigma_i^2}$ possui distribuição gama com $E(\tau_i) = 1$ e $Var(\tau_i) = 100$, $i = 1, 2$.

A inferência para esses modelos hierárquicos foi realizada utilizando algoritmo MCMC, obtendo-se amostras das distribuições a posteriori dos parâmetros. Este método foi implementado no *software* `WinBUGS`.

A convergência das cadeias foi monitorada graficamente por meio da análise das trajetórias de duas cadeias partindo de valores iniciais distintos. A fim de reduzir a autocorrelação existente entre os valores amostrados, foram incluídas na amostra da posteriori apenas 1 (um) valor a cada 100 iterações. A partir disto, concluiu-se que a convergência havia sido obtida após amostrarmos 150 e 100 valores, para os modelos utilizando os dados dos homens e mulheres, respectivamente.

Para compor a amostra, observou-se as cadeias por mais 1000 iterações, totalizando duas amostras de tamanho 1000 para cada parâmetro.

Os resultados obtidos a partir da utilização das quatro diferentes prioris foram analisados e comparados de acordo com os valores do DIC. A tabela 6.3 apresenta o DIC obtido para cada modelo para o sexo masculino.

Tabela 6.3: Comparação dos modelos hierárquicos com diferentes prioris por meio do DIC para o sexo masculino.

Priori	$\bar{D}$	$\hat{D}$	p_D	DIC
I	11746.400	11701.000	45.383	11791.800
II	11746.300	11701.700	44.584	**11790.900**
III	11746.900	11700.900	46.000	11792.900
IV	11746.600	11701.800	44.785	11791.400

As figuras 6.19 e 6.20 apresentam as médias a posteriori e respectivos intervalos de 95% de credibilidade dos parâmetros α_t e β_t para os homens, considerando a priori II, a qual apresenta o menor DIC.

Adicionalmente, a tabela 6.4 apresenta um sumário das distribuições a posteriori dos demais parâmetros do modelo considerando a priori II.

Tabela 6.4: Sumário da distribuição a posteriori dos parâmetros do modelo hierárquico com priori II para o sexo masculino.

Parâmetro	Média	Desvio Padrão	IC(95%)
a_0	-10.37	0.08531	(-10.53;-10.19)
a_1	-0.03795	0.006258	(-0.05052;-0.0252)
b_0	0.08492	0.006705	(0.07151;0.09786)
τ_1	29.7	10.2	(13.95;52.96)
τ_2	1088.0	323.1	(532.6;1793.0)

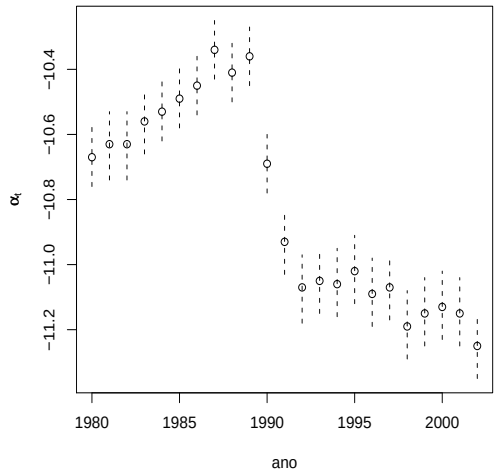

Figura 6.19: Médias a posteriori e respectivos intervalos de 95% de credibilidade de α_t para o modelo hierárquico com priori II para o sexo masculino.

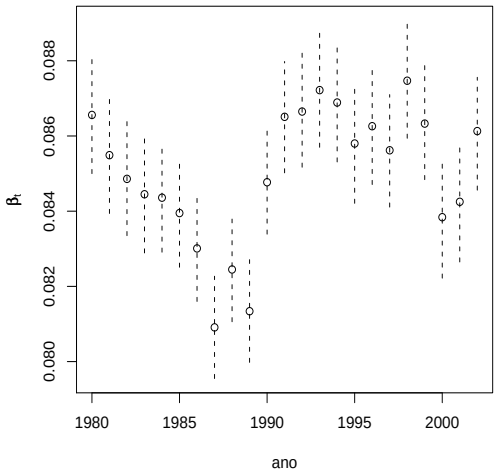

Figura 6.20: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_t para o modelo hierárquico com priori II para o sexo masculino.

A tabela 6.5 apresenta o DIC para os diferentes modelos hierárquicos considerados, para o sexo feminino.

Tabela 6.5: Comparação dos modelos hierárquicos com diferentes prioris por meio do DIC para o sexo feminino.

Priori	$\bar{D}$	$\hat{D}$	p_D	DIC
I	3551.720	3506.530	45.198	3596.920
II	3551.530	3507.400	44.135	3595.670
III	3551.430	3506.580	44.851	3596.280
IV	3551.460	3507.550	43.911	3595.370

Na análise dos dados das mulheres, de acordo com o critério do DIC, o melhor modelo é aquele que utiliza a priori IV. Entretanto, analisando as estimativas dos parâmetros deste modelo, observou-se que o parâmetro b_1 possui distribuição a posteriori centrada no valor zero, conforme mostra a figura 6.21. Uma vez que este parâmetro está associado ao efeito temporal, sua exclusão implica em um modelo com estrutura equivalente àquela fornecida pela priori II. Assim, considerou-se este último como a melhor escolha para a ánalise.

As figuras 6.22 e 6.23 apresentam as médias a posteriori e os respectivos intervalos de 95% de credibilidade para os parâmetros α_t e β_t, para as mulheres.

Adicionalmente, a tabela 6.6 apresenta o sumário da distribuição a posteriori dos demais parâmetros do modelo considerado.

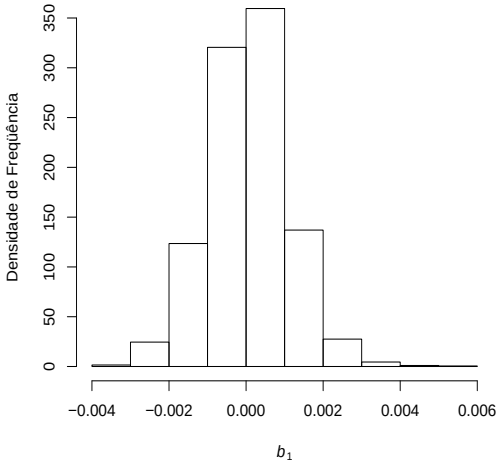

Figura 6.21: Histograma da distribuição a posteriori do parâmetro b_1 para o modelo hierárquico com priori IV para o sexo feminino.

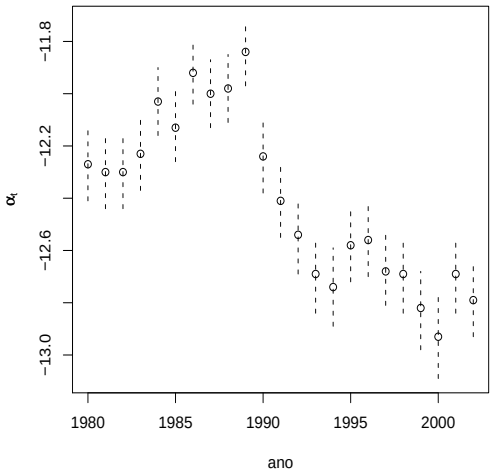

Figura 6.22: Médias a posteriori e respectivos intervalos de 95% de credibilidade de α_t para o modelo hierárquico com priori II para o sexo feminino.

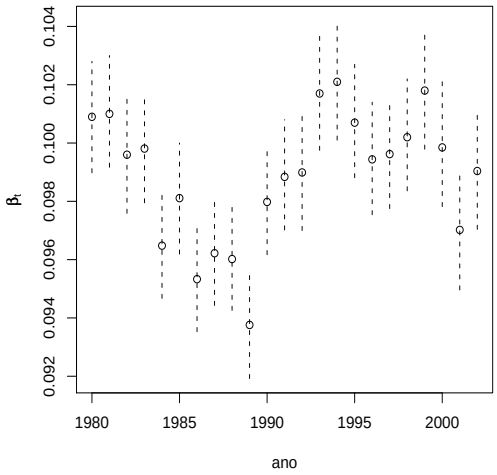

Figura 6.23: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_t para o modelo hierárquico com priori II para o sexo feminino.

Comparando a evolução das estimativas dos parâmetros α_t e β_t, podemos observar que os seus comportamentos são semelhantes àqueles obtidos nas ánalises utilizando o modelo dinâmico global, para ambos os sexos.

Comparando os valores do DIC das diferentes abordagens realizadas até aqui, notamos que a modelagem hierárquica forneceu melhores resultados do que àqueles obtidos pelo modelo dinâmico global, para ambos os sexos. Entretanto, estes valores do DIC ainda são superiores aos obtidos ao utilizar-se o modelo dinâmico local.

Tabela 6.6: Sumário da distribuição a posteriori dos parâmetros do modelo hierárquico com priori II para o sexo feminino.

Parâmetro	Média	Desvio Padrão	IC(95%)
a_0	-11.96	0.1022	(-12.16;-11.75)
a_1	-0.03729	0.007525	(-0.05218;-0.02292)
b_0	0.09897	0.006618	(0.08622;0.1121)
τ_1	20.82	7.289	(9.4;37.51)
τ_2	1091.0	334.0	(549.0;1840.0)

6.5 Modelo Linear Generalizado com Fator de Redução de Mortalidade

O modelo linear generalizado com fator de redução de mortalidade, conforme apresentado na seção 5.5, foi ajustado aos dados utilizando as prioris descritas abaixo:

- $\beta_x \sim$ Normal(0,1000), $\sum_{x=1}^{13} \beta_x = 1$;
- $\beta'_x \sim$ Normal(0,1000), $\sum_{x=1}^{13} \beta'_x = 1$;
- $\log(\phi) \sim$ Normal(0,1000).

E para os parâmetros γ_t foram utilizadas duas prioris distintas:

- $\gamma_t \sim$ Normal(0,1000), $\forall t$;
- $\gamma_t \sim$ Normal(γ_{t-1},1000), $t = 2, \ldots, 22$ com $\gamma_1 \sim$ Normal(0,1000).

Considerou-se o ano 2000 como padrão (t_0), uma vez que neste ano foi realizado o último censo no Brasil.

A inferência do modelo foi realizada utilizando-se o método de Monte Carlo via Cadeias de Markov (MCMC), implementado através do *software* `WinBUGS`.

A análise da convergência foi realizada por meio da monitoração visual de duas cadeias geradas a partir de diferentes valores iniciais dos parâmetros do modelo. A fim de eliminar a autocorrelação presente, considerou-se apenas 1 valor a cada 100 iterações para a realização da inferência. A convergência foi evidenciada após 100 iterações das cadeias.

Para a composição da amostra, observou-se a cadeia por outras 1.000 iterações, totalizando duas amostras de tamanho 1.000 para a análise da distribuição a posteriori de cada parâmetro do modelo.

Os diferentes modelos considerando diferentes anos de quebra t_q, conforme apresentados na seção 5.5.2, foram ajustados e comparados de acordo com o critério do DIC. A Tabela 6.7 apresenta os valores do DIC para os diferentes modelos considerados utilizando-se a tábua de mortalidade dos homens.

Analisando-se a Tabela 6.7 observa-se que o modelo com tempo de quebra $t_q = 1987$ e estrutura dinâmica no parâmetro γ apresentou o menor valor do DIC.

As figuras 6.24 e 6.25 apresentam as médias a posteriori e os respectivos intervalos de 95% de credibilidade para β_x e β'_x para este modelo. Analisando estas figuras, observa-se que os valores reduzidos de β_x para as faixas etárias entre 25 e 39 anos são compensadas por valores elevados de β'_x. Esta observação sugere que as reduções ocorridas nestas faixas etárias foram mais acentuadas que as demais até meados dos anos 80, mais especificamente, até o ano referente ao tempo de quebra. Esta redução diferenciada foi mais facilmente captada graças à inclusão do tempo de quebra no modelo, o qual possibilitou uma maior flexibilidade ao ajuste.

Adicionalmente, a figura 6.26 apresenta a evolução das médias a posteriori e respectivos intervalos de 95% de credibilidade para γ_t. Analisando as médias a posteriori e os respectivos intervalos de credibilidade de 95% de γ_t, para o sexo masculino, evidenciam-se reduções aproximadamente constantes deste efeito nas mortalidades ocorridas até 1997, seguidas por uma

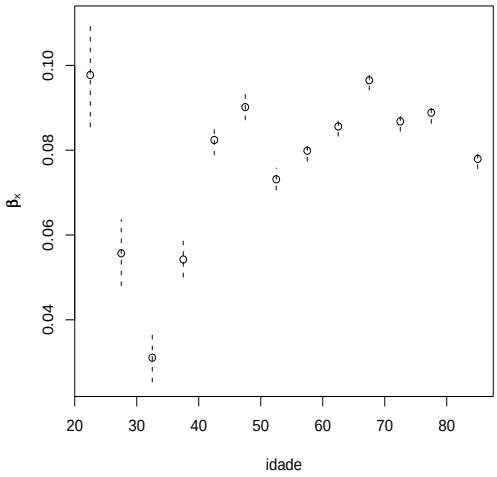

Figura 6.24: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_x para o modelo linear generalizado com fator de redução de mortalidade com $t_q = 1987$ e estrutura dinâmica no parâmetro γ para o sexo masculino.

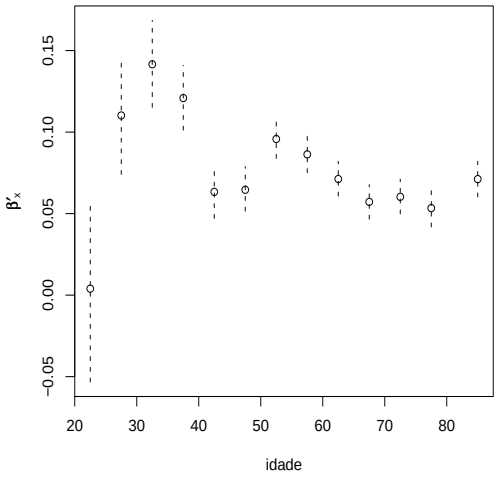

Figura 6.25: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β'_x para o modelo linear generalizado com fator de redução de mortalidade com $t_q = 1987$ e estrutura dinâmica no parâmetro γ para o sexo masculino.

Tabela 6.7: Valores do DIC considerando os diferentes pontos de quebra t_j para o sexo masculino.

t_j	s/ϕ	c/ϕ	γ dinâmico	γ dinâmico c/ϕ
–	2930.440	2932.200	2929.960	2932.970
1981	2913,100	2914.590	2913.740	2915.720
1982	2887.870	2890.610	2888.350	2891.000
1983	2866.460	2868.760	2867.060	2868.460
1984	2849.580	2852.610	2849.770	2851.800
1985	2843.610	2846.570	2844.230	2845.910
1986	2840.460	2843.390	2840.820	2843.170
1987	2840.510	2842.720	**2840.170**	2843.120
1988	2849.010	2850.830	2849.010	2850.830
1989	2858.870	2860.180	2858.870	2860.180
1990	2868.040	2869.200	2868.040	2869.200
1991	2869.690	2871.110	2869.690	2871.110
1992	2869.060	2870.640	2869.060	2870.640
1993	2865.350	2868.510	2865.350	2868.510
1994	2863.820	2865.760	2863.820	2865.760
1995	2861.030	2863.000	2861.030	2863.000
1996	2856.780	2859.420	2856.780	2859.420
1997	2857.160	2857.780	2857.160	2857.780
1998	2870.920	2872.060	2870.920	2872.060
1999	2914.580	2917.300	2914.580	2917.300

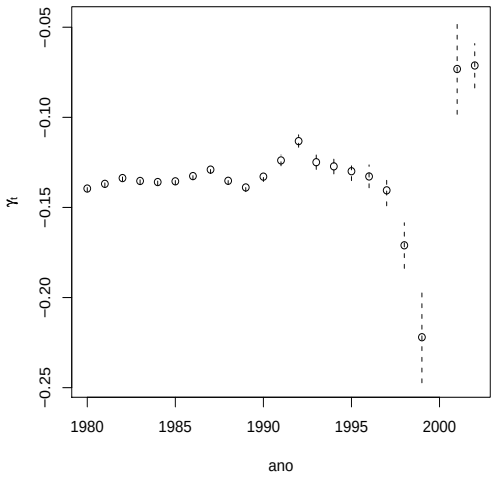

Figura 6.26: Médias a posteriori e respectivos intervalos de 95% de credibilidade de γ_t para o modelo linear generalizado com fator de redução de mortalidade para o modelo considerado para o sexo masculino.

amenização deste decréscimo a partir de 1998.

As estimativas dos fatores de redução de mortalidade, na escala logarítmica, para os homens, obtidos a partir do ajuste do modelo considerado, são apresentados na figura 6.27.

Analisando os fatores de redução de mortalidade estimados, pode-se observar uma redução bastante acentuada para as faixas etárias mais jovens (25 a 39 anos) em relação aos demais.

Os valores do DIC para os diferentes modelos considerando diferentes anos de quebra t_q, utilizando-se a tábua de mortalidade das mulheres são apresentados na Tabela 6.8.

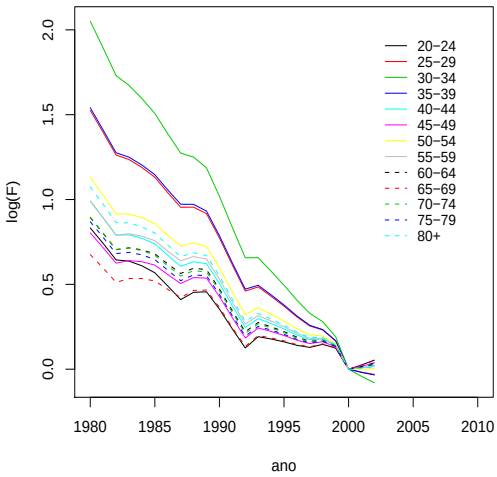

Figura 6.27: Médias a posteriori do fator de redução de mortalidade na escala logarítmica por faixa etária no período de 1980 a 2002 para o modelo considerado para o sexo masculino.

A partir da análise da tabela observa-se que o modelo com tempo de quebra $t_q = 1986$ e cuja distribuição a priori do parâmetro γ não apresenta estrutura dinâmica, apresentou o menor valor do DIC.

As figuras 6.28, 6.29 e 6.30 apresentam as médias a posteriori e respectivos intervalos de 95% de credibilidade dos parâmetros β_x, β'_x e γ_t, para o modelo considerado, para o sexo feminino. Analisando as médias a posteriori e os respectivos intervalos de credibilidade de 95% de β_x e β'_x, de forma semelhante ao ocorrido para o sexo masculino, conclui-se que a redução das mortalidades ocorridas na faixa etária de 25 a 29 anos até meados dos anos 80 justifica a inclusão do tempo de quebra no modelo. Da mesma forma, analisando-se as médias a posteriori e os respectivos intervalos de credibilidade de 95% de γ_t, para o sexo feminino, fica evidenciado o mesmo padrão

Tabela 6.8: Valores do DIC considerando os diferentes pontos de quebra t_j para o sexo feminino.

t_j	s/ϕ	c/ϕ	γ dinâmico	γ dinâmico c/ϕ
–	2754.590	2756.770	2755.100	2756.410
1981	2726.840	2729.140	2726.820	2728.910
1982	2697.040	2699.380	2696.980	2698.480
1983	2670.170	2673.030	2670.710	2673.130
1984	2643.380	2646.390	2643.700	2645.780
1985	2629.140	2630.970	2628.940	2630.280
1986	**2612.680**	2616.500	2613.780	2614.690
1987	2612.790	2614.690	2612.960	2614.330
1988	2613.520	2615.610	2613.520	2615.610
1989	2622.700	2624.650	2622.700	2624.650
1990	2639.190	2640.620	2639.190	2640.620
1991	2655.610	2656.020	2655.610	2656.020
1992	2666.180	2667.680	2666.180	2667.680
1993	2675.370	2677.990	2675.370	2677.990
1994	2682.170	2685.910	2682.170	2685.910
1995	2693.930	2694.280	2693.930	2694.280
1996	2680.130	2682.110	2680.130	2682.110
1997	2681.130	2683.470	2681.130	2683.470
1998	2689.110	2692.600	2689.110	2692.600
1999	2716.730	2721.640	2716.730	2721.640

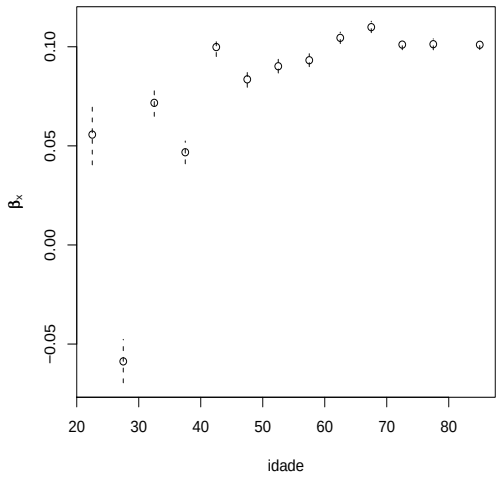

Figura 6.28: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β_x para o modelo linear generalizado com fator de redução de mortalidade com $t_q = 1986$ e sem estrutura dinâmica no parâmetro γ para o sexo feminino.

de redução na mortalidade verificado para o sexo masculino.

As estimativas dos fatores de redução de mortalidade, na escala logarítmica, para as mulheres, são apresentadas na figura 6.31.

Analisando o resultado para as diferentes faixas etárias, pode-se observar uma enorme diminuição do fator de redução de mortalidade das mulheres na faixa etária entre 25 e 29 anos, evidenciada pelo valor de $F(x,t)$ em 1980, o qual gira em torno de 54 vezes o fator do ano padrão.

Os efeitos de sobredispersão não se mostraram importantes nas análises de acordo com o criério do DIC, apresentando valores mais elevados em todos os modelos considerados.

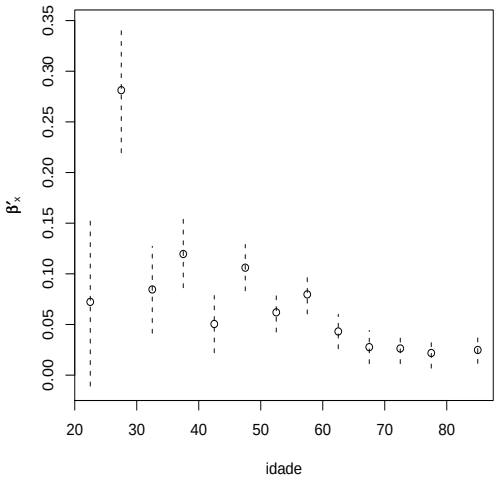

Figura 6.29: Médias a posteriori e respectivos intervalos de 95% de credibilidade de β'_x para o modelo linear generalizado com fator de redução de mortalidade com $t_q = 1986$ e sem estrutura dinâmica no parâmetro γ para o sexo feminino.

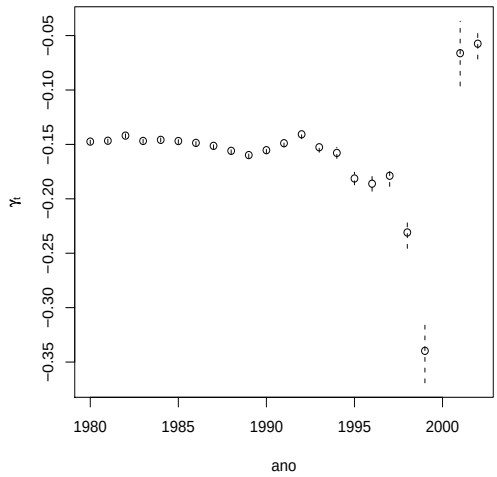

Figura 6.30: Médias a posteriori e respectivos intervalos de 95% de credibilidade de γ_t para o modelo linear generalizado com fator de redução de mortalidade para o modelo considerado para o sexo feminino.

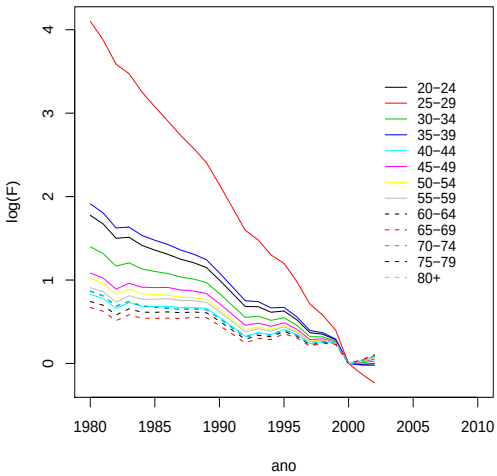

Figura 6.31: Médias a posteriori do fator de redução de mortalidade na escala logarítmica por faixa etária no período de 1980 a 2002 para o modelo considerado para o sexo feminino.

É interessante notar que os melhores modelos segundo o critério do DIC, para ambos os sexos, apontam para tempos de quebra anteriores a 1990. Apesar de ser apenas um indício, esta observação pode indicar que o aumento das mortes por causas mal definidas, ocorrida no início dos anos 90 com a publicação da portaria estadual descrita na seção 1.3, pode não ter sido de grande influência na redução acentuada das mortalidades por DISC.

Ainda segundo este critério, os modelos para ambos os sexos apresentaram os melhores resultados em comparação com o desempenho dos demais modelos considerados neste trabalho.

Capítulo 7

Considerações Finais

A análise dos dados de mortalidade por doenças isquêmicas do coração no estado do Rio de Janeiro no período de 1980 a 2002 evidenciou a tendência de redução das taxas observadas, assim como ocorre nos países desenvolvidos e em outras capitais e regiões brasileiras.

Os resultados também estiveram alinhados àqueles encontrados em trabalhos que utilizaram o mesmo banco de dados (Tura, 2005).

A análise dos dados foi realizada a partir da comparação de diferentes modelos e de uma grande quantidade de estruturas associadas a estes.

A utilização de uma abordagem bayesiana permitiu incorporar as incertezas associadas aos parâmetros a priori, as quais combinadas com as informações dos dados, forneceram as estimativas dos parâmetros de cada modelo ajustado.

Os modelos baseados no fator de redução de mortalidade apresentaram os melhores resultados em comparação com todos os modelos considerados

neste trabalho, para ambos os sexos.

Entre as vantagens observadas deste modelo, pode-se destacar a possibilidade de analisar a tendência da redução de mortalidade para as diferentes faixas etárias de forma conjunta. Esta característica evidencia a superioridade deste em relação aos modelos que ajustam curvas independentes para cada faixa etária por meio do ajuste de regressões lineares.

Ainda com relação ao resultado deste modelo, pode-se observar a vantagem da utilização do tempo de quebra, o qual permitiu um melhor ajuste dos fatores de redução aos dados de ambos os sexos. Como sugestão para trabalhos futuros, a consideração do tempo de quebra como um parâmetro a ser estimado pode facilitar o trabalho do pesquisador, deixando a tarefa de decidir o ano mais apropriado com os dados observados.

Adicionalmente, a utilização dos conceitos de modelos lineares generalizados dinâmicos (West *et al.*, 1985), por meio da introdução de estruturas dinâmicas em alguns parâmetros dos modelos forneceu bons resultados em todos os modelos considerados.

Por fim, a realização deste trabalho proporcionou o estudo de um tema de grande importância por meio da avaliação dos resultados obtidos utilizando uma grande quantitade de modelos, sob vários enfoques.

Referências Bibliográficas

[1] Armazém de Dados, Prefeitura da Cidade do Rio de Janeiro. GIE/COE/SSC/SMS-RJ. Disponível em: http://www.saude.rio.rj.gov.br, jul. 2006.

[2] Cordeiro, G.M. e Lima Neto, E.A. (2004). *Modelos Paramétricos*. 16° SINAPE.

[3] Daniel, E., Germiniani, H., Nazareno, E. R., Braga, S. V., Winkler, A. M. e Cunha, C. L. P. (2005). Tendência da Mortalidade por Doenças Isquêmicas do Coração na Cidade de Curitiba - Brasil, de 1980 a 1998. *Arquivos Brasileiros de Cardiologia*, vol 85, n° 2, ago.

[4] DeGroot, M.H. e Schervish, M.J. (2002). *Probability and Statistics* 3rd ed. Addison-Wesley.

[5] Ferreira, M.A.R. e Gamerman, D. (1998). Análise Bayesiana de Séries Epidemiológicas de Contagem via Modelos Dinâmicos Bayesianos Generalizados. *Cadernos de Saúde Coletiva*, 6(2).

[6] Gamerman, D. e Migon, H. (1993). Dynamic Hierarquical Models. *Journal of the Royal Statistical Society*, B, 55, 629-642.

[7] Gamerman, D. (1997). *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference*. Chapman and Hall.

[8] Gelfand, A.E. e Smith, A.F.M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409.

[9] Gelfand, A.E. e Ghosh, S.K. (1998). Model Choice: A Minimun Posterior Predictive Loss Approach. *Biometrika*, 85, 1-11.

[10] Gelman, A. e Rubin, D.B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-472.

[11] Geman, S. e Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

[12] Geweke, J. (1992). Evaluating the Accuracy of Sampling-Based Approaches to Calculating Posterior Moments. Em *Bayesian Statistics* 4, 169-193 (Editors Bernardo, J.M., Berger, J.O., Dawid, A.P. e Smith, A.F.M.).

[13] Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*, Vol. 115, 513-585.

[14] Hastings, W.K. (1970). Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57, 97-109.

[15] IBGE. Consulta eletrônica. Disponível em: http://www.ibge.gov.br/

[16] Laurenti, R., Gotlieb, S. L., Souza, J. M. P., Fonseca, L. A. M. e Jorge, M. H. P. (1981). Características da Mortalidade por Doença Isquêmica do Coração em Adultos de 15 a 74 anos no município de São Paulo. *Arquivos Brasileiros de Cardiologia*, vol.36, nº 2, 85-89.

[17] Laurenti, R. (1982). Epidemiologia das Doenças Cardiovasculares no Brasil. ıArquivos Brasileiros de Cardiologia, vol.38, nº 4, 243-248.

[18] Laurenti, R., Buchalla, C. M. e Caratin, C. V. S. (2000). Doença Isquêmica do Coração. Internações, Tempo de Permanência e Gastos. Brasil, 1993 a 1997. *Arquivos Brasileiros de Cardiologia*, vol.74, nº 6, 483-487.

[19] Lee, R.D. e Carter, L. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, 87, 659-671.

[20] Lolio, C. A. e Laurenti, R. (1986). Mortalidade por Doença Isquêmica do Coração no Município de São Paulo. Evolução de 1950 a 1981 e Mudanças Recentes na Tendência. *Arquivos Brasileiros de Cardiologia*, vol.46, nº 3, 153-156.

[21] Lolio, C. A., Lotufo, P. A., Lira, A. C., Zanetta, D. M. T. e Massad, E. (1994). Tendência da Mortalidade por Doença Isquêmica do Coração nas Capitais de Regiões Metropolitanas do Brasil, 1979-89. *Arquivos Brasileiros de Cardiologia*, vol.64, nº 3, 213-216.

[22] Lotufo, P. A. e Lolio, C. A. (1993). Tendência da Mortalidade por Doença Isquêmica do Coração no Estado de São Paulo: 1970 a 1989. *Arquivos Brasileiros de Cardiologia*, vol.61, nº 3, 149-153.

[23] Mansur, A. P., Favarato, D., Souza, M. F. M., Avakian, S. D., Aldrighi, J. M., César, L. A. M. e Ramires, J. A. F. (2001). Tendência do Risco

de Morte por Doenças Circulatórias no Brasil de 1979 a 1996. *Arquivos Brasileiros de Cardiologia*, vol.76, nº 6, 497-503.

[24] Marcopito, L. F. (2003). Doenças isquêmicas do coração: mortalidade em nativos e migrantes, São Paulo, 1979-1998. *Revista de Saúde Pública*, vol. 37, nº 6, dez.

[25] McCullagh, P. e Nelder, J.A. (1989). *Generalized Linear Models* 2nd ed. Chapman and Hall, Londres.

[26] Metropolis, N., Rosenbulth, A.W., Rosenbulth, M.N., Teller, A.H. e Teller, E. (1953). Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics*, 21, 1089-1091.

[27] Migon, H.S. e Gamerman, D. (1993). Generalized Exponential Growth Models - a Bayesian Approach. *Journal of Forecasting*, 12, 573-584.

[28] Migon, H.S. e Gamerman, D. (1999). *Statistical Inference - an integrated approach*. Arnold.

[29] Ministério da Saúde - DATASUS. Taxa de Mortalidade por Doenças Isquêmicas do Coração. Ficha de qualificação - Indicadores de Mortalidade. Disponível em: http://www.datasus.gov.br.

[30] Moraes, S. A., Rezende, M. H. V. e Freitas, I. C. M. (2000). Tendência de Mortalidade por Doença Isquêmica do Coração no Município de Goiânia - Brasil na Série Histórica entre 1980 e 1994. *Arquivos Brasileiros de Cardiologia*, vol.74, nº 6, 493-497.

[31] Neves, C.R. e Migon, H.S. (2007). Bayesian Graduation of Mortality Rates: An Application to Reserve Evaluation. *Insurance: Mathematics and Economics*, vol 40, 424-434.

[32] Oliveira, G.M.M, Klein, C.H. e Silva, N.A.S. (2003). Análise Crítica das Mudanças das Taxas de Mortalidade por Doenças do Aparelho Circulatório Ocorridas entre 1980 e 2000, no Estado do Rio de Janeiro. *Revista da SOCERJ*, 16(2), 95-100.

[33] Oliveira, G. M. M., Klein, C. H. e Silva, N. A. S. (2005)[a]. Efeitos de Idade, Geração e Período na Mortalidade por Doenças Isquêmicas do Coração e por Causas Mal Definidas nos Estados do Rio de Janeiro, São Paulo e Rio Grande do Sul nos Anos de 1980, 1990 e 2000 - Brasil. *Revista da SOCERJ*, vol 18, nº 2, mar-abr.

[34] Oliveira, G. M. M., Klein, C. H. e Silva, N. A. S. (2005)[b]. Efeitos de Idade, Geração e Período na Mortalidade por Doenças Cerebrovasculares e por Todas as Causas nos Estados do Rio de Janeiro, São Paulo e Rio Grande do Sul nos Anos de 1980, 1990 e 2000 - Brasil. *Revista da SOCERJ*, vol 18, nº 4, jul-ago.

[35] Oliveira, G. M. M., Klein, C. H. e Silva, N. A. S. (2005)[c]. Mortalidade por Doenças Isquêmicas do Coração, Doenças Cerebrovasculares e Causas Mal Definidas nas Regiões de Saúde do Estado do Rio de Janeiro, no período de 1980 a 2000. *Revista da SOCERJ*, vol 18, nº 1, jan-fev, 13-22.

[36] Oliveira, G. M. M., Silva, N. A. S. e Klein, C. H. (2005)[d]. Mortalidade compensada por doenças cardiovasculares no período de 1980 a 1999 - Brasil. *Arquivos Brasileiros de Cardiologia*, vol.85, nº 5, nov.

[37] Oliveira, G. M. M., Klein, C. H. e Silva, N. A. S. (2006). Mortalidade por doenças cardiovasculares em três Estados do Brasil de 1980 a 2002. *Revista Panamericana de Salud Pública*, vol.19, nº 2, 85-93.

[38] Organização Mundial de Saúde - Manual da Classificação Internacional de Doenças, Lesões e Causas de Óbitos, 9ª Revisão, 1975. Centro da OMS para Classificação das Doenças em Português. São Paulo; 1978.

[39] Organização Mundial de Saúde. Classificação Estatística Internacional de Doenças e Problemas relacionados à Saúde: Classificação Internacional de Doenças. 10ª Revisão. São Paulo: EDUSP; 1995:1.

[40] Paes, N. A. (1985). Mortalidade em Recife: aplicação de um modelo de riscos competitivos. *Revista de Saúde Pública*, vol.19, nº 3, jun.

[41] Passos, L. C. S., Lopes, A. A., Lessa, I., Sanches, A. e Santos-Jesus, R. (2000). Tendência da Mortalidade por Infarto Agudo do Miocárdio (1981 a 1996) na Cidade de Salvador, Brasil. ıArquivos Brasileiros de Cardiologia, vol.74, nº 4, 329-331.

[42] Portaria número 550 de 23 de janeiro de 1990, publicada no Diário Oficial de 29 de janeiro de 1990.

[43] Renshaw, A.E.e Haberman, S. (2000). Modelling for Mortality Reduction Factors. *Actuarial Research Paper No. 127*, City University, Londres.

[44] Renshaw, A.E.e Haberman, S. (2003). Lee-Carter Mortality Forecasting: a Parallel Generalized Linear Modelling Approach for England and Wales Mortality Projections. *Applied Statistics*, 52, 119-137.

[45] Renshaw, A.E.e Haberman, S. (2003). On the Forecasting of Mortality Reduction Factors. *Insurance: Mathematics and Economics*, 32, 379-401.

[46] Renshaw, A.E.e Haberman, S. (2003). Lee-Carter Mortality Forecasting with Age-specific Enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.

[47] Secretaria de Vigilância em Saúde - Ministério da Saúde. (2004). Inquérito domiciliar sobre comportamentos de risco e morbidade referida de doenças e agravos não transmissíveis: Brasil, 15 Capitais e Distrito Federal 2002-2003. INCA. Rio de Janeiro.

[48] Sistema de Informações sobre Mortalidade. Consulta eletrônica. Disponível em: http://www.datasus.gov.br/catalogo/sim.htm

[49] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. e Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, B, 64, 583-639.

[50] Thomas, A., Spiegelhalter, D.J. e Gilks, W.R. (1992). BUGS: A program to perform Bayesian inference using Gibbs Sampling. . Em *Bayesian Statistics* 4, 837-842 (Editors Bernardo, J.M., Berger, J.O., Dawid, A.P. e Smith, A.F.M.).

[51] Tomassini, H. C. B., Alves, M. G. M., Claro, L. B. L., Pacheco, A. G. F. e Almeida, M. T. C. G. N. (2003). Evolução da população e da mortalidade em Niterói, Rio de Janeiro, Brasil, 1979 a 1997. *Cadernos de Saúde Pública*, vol.19, nº 6, nov-dez.

[52] Tuljapurkar, S., Li, N. e Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Letters to Nature*, vol 405, jun.

[53] Tura, B.R. (2005). *Análise da Dinâmica de Mortalidade das Doenças do Aparelho Cardiovascular Utilizando a Equação de Gompertz*, tese de doutorado, UFRJ, Rio de Janeiro.

[54] Tura, B. R. T., Silva, N. A. S. e Pereira, B. B. (2006). Associação entre Renda Per Capita e Mortalidade por Doença Cardiovascular. *Revista da SOCERJ*, vol 19, nº 3, mai/jun.

[55] West, M. e Harrison, P.J. (1997). *Bayesian Forecasting and Dynamic Models* 2nd ed. Springer Verlag, New York.

[56] West, M. e Harrison, P.J. and Migon, H.S. (1985). Dynamic generalized linear models and Bayesian forecasting (com discussão). *Journal of the American Statistical Association*, 80, 73-96.

Apêndice

Deviance Information Criterion (DIC)

Um modo natural de comparar modelos é usar critérios baseados em medidas de ajuste dos dados aos diferentes modelos respeitando o princípio da parcimônia, isto é, escolher entre modelos que expliquem consideravelmente bem um certo conjunto de dados o que tiver o menor número de parâmetros.

O *Deviance Information Criterion* (DIC) (Spiegelhalter *et al.*, 2002) é um critério bayesiano de seleção de modelos onde serão computadas as seguintes estatísticas:

- $\bar{D}$: Representa a média a posteriori da *deviance*. A *deviance* é definida como

  $$D = -2\ln(p(\mathbf{y}|\boldsymbol{\theta})),$$

  onde $p(\mathbf{y}|\boldsymbol{\theta})$ representa a função de verossimilhança.

- $\hat{D}$: É uma estimativa pontual da *deviance* obtida a partir da utilização da média a posteriori de $\boldsymbol{\theta}$, denotada por $\bar{\boldsymbol{\theta}}$, na expressão da *deviance*. Logo,

  $$\hat{D} = -2\ln(p(\mathbf{y}|\bar{\boldsymbol{\theta}})).$$

  Dependendo da parametrização utilizada na especificação das prioris

do modelo o valor de $\hat{D}$ pode ser alterado.

- p_D: Representa o número efetivo de parâmetros e é dada por $p_D = \bar{D} - \hat{D}$. Para modelos não-hierárquicos, o valor de p_D pode ser aproximadamente igual ao verdadeiro número de parâmetros.

A partir destas estatísticas, o DIC é definido como:

$$DIC = \bar{D} + p_D = \hat{D} + 2p_D.$$

O cálculo do DIC vem incorporado na versão 1.4.1 do `WinBUGS`. O modelo que apresenta o menor DIC é considerado como aquele que melhor pode predizer um novo conjunto de dados com a mesma estrutura dos dados observados.

O critério DIC assume que a média a posteriori é uma boa estimativa dos parâmetros do modelo. Assim, nos casos em que estas distribuições são multimodais ou nos casos em que há uma acentuada assimetria, o DIC não é recomendado.

O DIC é semelhante ao critério clássico de comparação de modelos AIC (Akaike's Information Criterion), mas difere substancialmente do critério BIC e do fator de Bayes.

Livros Grátis

(http://www.livrosgratis.com.br)

Milhares de Livros para Download:

Baixar livros de Administração
Baixar livros de Agronomia
Baixar livros de Arquitetura
Baixar livros de Artes
Baixar livros de Astronomia
Baixar livros de Biologia Geral
Baixar livros de Ciência da Computação
Baixar livros de Ciência da Informação
Baixar livros de Ciência Política
Baixar livros de Ciências da Saúde
Baixar livros de Comunicação
Baixar livros do Conselho Nacional de Educação - CNE
Baixar livros de Defesa civil
Baixar livros de Direito
Baixar livros de Direitos humanos
Baixar livros de Economia
Baixar livros de Economia Doméstica
Baixar livros de Educação
Baixar livros de Educação - Trânsito
Baixar livros de Educação Física
Baixar livros de Engenharia Aeroespacial
Baixar livros de Farmácia
Baixar livros de Filosofia
Baixar livros de Física
Baixar livros de Geociências
Baixar livros de Geografia
Baixar livros de História
Baixar livros de Línguas

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo