

Universidade Federal do Rio de Janeiro

**Modelos espaço-temporais para órbitas por  
causas externas**

Luiz Ledo Mota Melo Junior

Rio de Janeiro

2008

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Luiz Ledo Mota Melo Junior

**Modelos espaço-temporais para órbitas por causas externas**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Orientadores:

Marina Silva Paez

Ricardo Sandes Ehlers

Departamento de Métodos Estatísticos  
Instituto de Matemática  
Universidade Federal do Rio de Janeiro

agosto de 2008

# Modelos espaço-temporais para óbitos por causas externas

**Luiz Ledo Mota Melo Junior**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Aprovada por:

Banca examinadora:

---

Marina Silva Paez

Instituto de Matemática - UFRJ

---

Fernando Antonio da Silva Moura

Departamento de Estatística - UFRJ

---

Ricardo Sandes Ehlers

Departamento de Estatística - UFPR



Para minha avó, Alcidia.

# Agradecimentos

Vou fazer os agradecimentos primeiramente à minha família. Devo tudo a minha avó. Minha mãe me ajudou muito, e minha tia Renata e meu primo João Alberto também contribuíram bastante, mesmo quando outros familiares não acreditaram em mim. Agradeço também ao meu irmão mais novo e espero estar sendo um bom exemplo para ele.

Aos amigos de Curitiba, agradeço principalmente ao Silvio, pois me apoiou muito para eu ir ao Rio de Janeiro.

Um agradecimento especial ao professor Ricardo Ehlers, pois sem ele não teria conseguido chegar a uma pós-graduação. Na UFPR aprendi muito com os professores Fernando Lucambio, Paulo Justiniano (considero ele meu terceiro orientador) e Joel Mauricio. Na UFRJ agradeço a Marina Paez, pois antes dela eu era um analfabeto estatístico.

No Rio de Janeiro, certamente sem a ajuda do Fernando Nascimento e do Vinicius Israel, não teria conseguido concluir o curso. Também ressalto a Valmária por me apoiar sempre.

Agradeço aos funcionários e professores da UFRJ pela experiência ótima que tive nessa instituição. Também agradeço ao CNPQ por ter disponibilizado uma bolsa de estudos.

# RESUMO

Em óbitos por causas externas, estão agrupados homicídios, suicídios e acidentes de trânsito. Essa variável será estudada nos 399 municípios do Estado do Paraná, nos anos de 1979 a 2004. O objetivo desse trabalho é estudar a dinâmica dessa variável no tempo e no espaço, para detectar as regiões críticas do Paraná e observar no espaço a formação de clusters. Propusemos uma modelagem onde o número de óbitos por causas externas ( $y_{it}$ ) tem distribuição de Poisson, com valor esperado dado pela multiplicação de dois componentes: um representando a densidade populacional ( $e_{it}$ ), e outro o risco relativo de cada município ( $\psi_{it}$ ). Diferentes modelagens foram propostas para o logaritmo do risco relativo. Algumas dessas propostas incluem uma estrutura espacial com vizinhança baseada em fronteira, outras assumem independência espacial. Em algumas propostas incorporamos como covariável o logaritmo do número de escolas de cada município dividido pelas suas respectivas populações. Por fim, modelamos conjuntamente as variáveis óbitos por causas externas e número de casos de tráfico de drogas para um período de tempo fixo (ano de 2004), testando estruturas com e sem dependência espacial.

# ABSTRACT

In this work we study deaths by unnatural causes in the 399 municipalities of Parana State, Brazil, from 1979 to 2004. This variable includes homicides, suicides and deaths by traffic accidents. Our aim is to study the dynamic of this variable in time and space, to detect the critic regions in Paraná and observe the formation of clusters in space. We propose a model where the number of deaths by unnatural causes ( $y_{it}$ ) is Poisson distributed, with its expectation given by the product of two components: one representing the pupulational density ( $e_{it}$ ), and the other representing the relative risk of each municipality ( $\psi_{it}$ ). Different models were proposed for the logarithm of the relative risk. Some of these proposals include a spatial structure with neighbourhood based in the borders of the municipalities, others assume spacial independency. In some proposals we incorporate as a covariable the logarithm of the number of schools in each municipality divided by their respective populations. Finally, we model the variables deaths by unnatural causes and number of cases of drug dealing jointly for a fixed period of time (the year of 2004), where we tested structures with and without spatial dependency.

# Sumário

<b>Introdução</b>	<b>1</b>
<b>1 Inferência Bayesiana</b>	<b>4</b>
1.1 Distribuição a Priori . . . . .	5
1.1.1 Prioris Conjugadas . . . . .	5
1.1.2 Prioris não Informativas . . . . .	5
1.1.3 Prioris Hierárquicas . . . . .	6
1.2 Estimação . . . . .	7
1.2.1 Estimação Pontual . . . . .	7
1.2.2 Estimação por Intervalos . . . . .	8
<b>2 Modelos Lineares</b>	<b>9</b>
2.1 Modelos Lineares Gaussianos . . . . .	9
2.2 Modelos Dinâmicos Lineares . . . . .	10
2.3 Modelos Lineares Generalizados . . . . .	13
2.4 Modelos Lineares Dinâmicos Generalizados . . . . .	14
<b>3 Modelos estruturados no espaço</b>	<b>16</b>
3.1 Campos Aleatórios Markovianos Gaussianos . . . . .	16
3.2 Modelos <i>CAR</i> Multivariados . . . . .	18
<b>4 Métodos Computacionalmente Intensivos</b>	<b>20</b>
4.1 Introdução . . . . .	20
4.2 Algoritmo Metropolis-Hastings . . . . .	20
4.3 Metropolis-Hastings com uma variável por vez . . . . .	21
4.4 Amostrador de Gibbs . . . . .	22
4.5 Diagnósticos de Convergência . . . . .	23

<b>5</b>	<b>Análise do Número de Óbitos por Causas Externas no estado do Paraná</b>	<b>25</b>
5.1	Aspectos Gerais . . . . .	26
5.2	Cálculo do valor esperado de óbitos por município . . . . .	35
5.3	Cálculo do Standardized Mortality Ratio (SMR) . . . . .	40
5.4	Relação entre óbitos por causas externas e tráfico de drogas . . . . .	42
<b>6</b>	<b>Modelagem Univariada</b>	<b>45</b>
6.1	Especificação dos Modelos . . . . .	45
6.2	Inferência . . . . .	48
6.3	Critério <i>DIC</i> (Deviance Information Criterion) para Comparação de Modelos . . . . .	51
6.4	Resultados Obtidos . . . . .	52
<b>7</b>	<b>Modelos Multivariados Espaciais</b>	<b>62</b>
7.1	Inferência . . . . .	63
7.2	Resultados Obtidos . . . . .	64
	<b>Conclusão</b>	<b>69</b>
	<b>Referências Bibliográficas</b>	<b>71</b>
	<b>Apêndice</b>	<b>74</b>

# Lista de Figuras

5.1	<i>Mapa do Estado do Paraná</i>	27
5.2	<i>Número médio de óbitos por causas externas versus Ano</i>	29
5.3	<i>Número de óbitos por causas externas das principais cidades do Estado do Paraná</i>	30
5.4	<i>Média de óbitos por causas externas nos 26 anos estudados</i>	31
5.5	<i>Média da covariável escola nos 26 anos estudados</i>	31
5.6	<i>Número de escolas / população - 1980</i>	32
5.7	<i>Número de escolas / população - 1990</i>	32
5.8	<i>Número de escolas / população - 1996</i>	33
5.9	<i>Número de escolas / população - 2004</i>	33
5.10	<i>Óbitos por causas externas no ano de 1980</i>	33
5.11	<i>Óbitos por causas externas no ano de 1990</i>	33
5.12	<i>Óbitos por causas externas no ano de 1996</i>	34
5.13	<i>Óbitos por causas externas no ano de 2004</i>	34
5.14	<i>Valor esperado de óbitos por causas externas em 1980</i>	35
5.15	<i>Valor esperado de óbitos por causas externas em 1990</i>	35
5.16	<i>Valor esperado de óbitos por causas externas em 1996</i>	36
5.17	<i>Valor esperado de óbitos por causas externas em 2004</i>	36
5.18	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1980</i>	38
5.19	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1990</i>	38
5.20	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1996</i>	38
5.21	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado em 2004</i>	38
5.22	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado com Curitiba em 2004</i>	39
5.23	<i>Gráfico óbitos por causas externas versus seu respectivo valor esperado sem Curitiba em 2004</i>	39
5.24	<i>Resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado com Curitiba 2004</i>	39

5.25	<i>Resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado sem Curitiba 2004</i> . . . . .	39
5.26	<i>Gráfico dos resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado com Curitiba versus logaritmo covariável escola padronizada em 2004.</i> . . . . .	40
5.27	<i>Gráfico dos resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado sem Curitiba versus logaritmo covariável escola padronizada em 2004.</i> . . . . .	40
5.28	<i>SMR óbitos por causas externas em 2004</i> . . . . .	41
5.29	<i>Gráfico óbitos por causas externas versus número de casos de tráfico de drogas em 2004.</i> . . . . .	42
5.30	<i>Gráfico óbitos por causas externas versus número de casos de tráfico de drogas em 2004 sem Curitiba.</i> . . . . .	42
5.31	<i>SMR número de casos de tráfico de drogas em 2004.</i> . . . . .	43
6.1	<i>Gráfico de máxima densidade a posteriori do Intercepto do M5 com P1 nos 26 anos estudados</i>	54
6.2	<i>Gráfico de máxima densidade a posteriori do Intercepto do M5 com P2 nos 26 anos estudados</i>	54
6.3	<i>Gráfico dos intervalos de máxima densidade a posteriori da precisão do M6 com P1 nos 26 anos estudados</i> . . . . .	55
6.4	<i>Gráfico dos intervalos de máxima densidade a posteriori da precisão do M6 com P2 nos 26 anos estudados</i> . . . . .	55
6.5	<i>Gráfico dos intervalos de máxima densidade a posteriori da precisão do modelo selecionado nos 26 anos estudados</i> . . . . .	56
6.6	<i>Gráfico dos intervalos de máxima densidade a posteriori do intercepto do modelo selecionado nos 26 anos estudados</i> . . . . .	56
6.7	<i>Gráfico SMR versus logaritmo da população em 1980</i> . . . . .	57
6.8	<i>Média a posteriori versus logaritmo da população em 1980</i> . . . . .	57
6.9	<i>SMR versus logaritmo da população em 1990</i> . . . . .	58
6.10	<i>Média a posteriori versus logaritmo da população em 1990</i> . . . . .	58
6.11	<i>SMR versus logaritmo da população 1996</i> . . . . .	59
6.12	<i>Média a posteriori versus logaritmo da população em 1996</i> . . . . .	59
6.13	<i>Gráfico SMR versus logaritmo da população 2004</i> . . . . .	59
6.14	<i>Média a posteriori versus logaritmo da população em 2004</i> . . . . .	59
6.15	<i>Média a posteriori do componente estocástico estruturado do modelo selecionado em 2004</i>	60
6.16	<i>Média a posteriori do risco relativo 1980 do modelo selecionado.</i> . . . . .	61
6.17	<i>Média a posteriori do risco relativo 2004 do modelo selecionado.</i> . . . . .	61
7.1	<i>Histograma da média a posteriori do intercepto da variável óbitos por causas externas</i> . . . . .	65



7.2	<i>Histograma da média a posteriori do intercepto da variável tráfico de drogas . . . . .</i>	65
7.3	<i>Histograma da média a posteriori da variância da variável óbitos por causas externas. . .</i>	66
7.4	<i>Histograma da média a posteriori da variância da variável tráfico de drogas. . . . .</i>	66
7.5	<i>Histograma da média a posteriori da correlação. . . . .</i>	67
7.6	<i>Média a posteriori do componente estocástico estruturado de óbitos por causas externas.</i>	68
7.7	<i>Média a posteriori do componente estocástico estruturado do número de casos de tráfico de drogas. . . . .</i>	68
7.8	<i>Média a posteriori do risco relativo do número de casos de tráfico de drogas . . . . .</i>	68
7.9	<i>Média a posteriori do risco relativo da variável óbitos por causas externas . . . . .</i>	68

# Lista de Tabelas

5.1	População das principais cidades do Paraná . . . . .	28
6.1	Tabela DIC dos 14 modelos univariados . . . . .	52
6.2	Média a Posteriori . . . . .	53
6.3	Desvio Padrão a Posteriori . . . . .	53
7.1	Resultados sob o Modelo 1 . . . . .	66

# Introdução

Com o interesse cada vez maior das autoridades públicas em tentar prever e prevenir epidemias de doenças como Dengue e AIDS, e problemas sociais como o da criminalidade, diversas metodologias estatísticas têm sido propostas para aplicações em estudos de epidemiologia, saúde e criminalidade. Modelos espaço-temporais têm sido largamente utilizados para estudar o comportamento das epidemias de doenças (ou criminalidade) em uma região de interesse, com a finalidade de fazer previsões e ajudar no controle das mesmas. Os modelos espaço-temporais podem ser aplicados a qualquer conjunto de dados coletados em diferentes localizações associadas a um ponto (quando sabemos a exata localização da ocorrência do evento ou medição) ou região do espaço (quando os dados são agregados em áreas disjuntas), observados em períodos de tempo diversos.

Na área de segurança pública, saúde e criminalidade, muitas vezes existe a necessidade de identificar quais regiões, dentro de uma determinada área geográfica, tem mais alto nível de risco. Em geral, utiliza-se para isso variáveis como número de óbitos por violência, acidentes e número de indivíduos afetados por uma particular doença. Uma forma de estudar o comportamento da variável de interesse é através do mapeamento do risco relativo, que pode nos indicar padrões espaciais que auxiliem as autoridades públicas a determinar os principais focos do problema.

Alguns trabalhos de epidemiologia utilizam a taxa de mortalidade padronizada, que é definida por  $SMR_{it} = \frac{y_{it}}{e_{it}}$ , em inglês é denominado Standardized Mortality Rate (SMR). Uma crítica para esta metodologia é superestimar os riscos relativos de pequenas populações (Mollié 1996). Um dos objetivos desse trabalho é propor uma metodologia Bayesiana para calcular o risco relativo e comparar com o  $SMR_{it}$ . Na abordagem

Bayesiana, o logaritmo do risco relativo é modelado através de uma regressão linear.

Modelos espaço-temporais Bayesianos também vem sendo largamente utilizados, com a vantagem de levar em consideração a incerteza associada aos parâmetros desconhecidos do modelo. Em Paez e Gamerman (2003) os autores propuseram modelos espaço-temporais para poluição ambiental coletada em estações de monitoramento na Região Metropolitana do Rio de Janeiro. Os modelos propostos supunham efeitos aditivos de espaço e tempo e diferentes combinações desses efeitos foram testadas para a escolha do melhor modelo. Ferreira e Schimdt (2006) propuseram modelos espaço-temporais para modelar o risco relativo de Dengue no Rio de Janeiro e Nobre, Schimdt e Lopes (2005) propuseram modelos espaço-temporais para modelar a incidência de Malária no Pará. Neste trabalho nosso interesse é estudar a dinâmica espacial e temporal da variável óbitos por causas externas no Estado do Paraná, variável esta que agrupa homicídios, suicídios e acidentes de trânsito. Os óbitos serão estudados nos 399 municípios do Estado do Paraná nos anos de 1979 a 2004. Primeiramente propomos um modelo univariado, onde o nosso objetivo é estudar a dinâmica da variável de interesse no tempo e no espaço. Para tentar entender como essa variável se comporta serão testados modelos com efeitos espaciais e determinísticos, e incluindo possíveis variáveis explicativas. Para modelar a estrutura espacial utilizamos Campos Aleatórios Markovianos Gaussianos (CAMG) (Besag et.al, 1991), e para modelar a estrutura temporal utilizamos Modelos Dinâmicos (West e Harrison, 1997).

Em uma segunda proposta, modelamos conjuntamente óbitos por causas externas e o número de casos de tráfico de drogas no Estado do Paraná. Essas duas variáveis estão correlacionadas devido à violência gerada com o tráfico. Para modelar a estrutura espacial conjunta será utilizado o modelo CAMG Multivariado, proposto no artigo de Ma e Carlin (2007), que é uma extensão do Modelo CAMG utilizado nos modelos univariados.

Um dos objetivos principais desse trabalho é estimar o risco relativo, mas a distribuição a posteriori desse parâmetro não possui forma analítica fechada e por isso são utilizados os métodos computacionais de Monte Carlo em cadeias de Markov (MCMC) usando os algoritmos de amostrados de Gibbs e Metropolis-Hastings, com o objetivo de obter uma amostra da distribuição de interesse.

Esse trabalho será dividido da seguinte forma: no capítulo 1 será feita uma revisão de inferência Bayesiana; no capítulo 2 serão abordados os modelos lineares Gaussianos; no capítulo 3 serão discutidos modelos estruturados no espaço; no capítulo 4 será feita uma revisão dos métodos computacionalmente intensivos MCMC; no capítulo 5 será realizado uma análise descritiva de óbitos por causas externas, número de casos de tráfico de drogas e da covariável escola; nos capítulos 6 e 7 serão propostos modelos espaço-temporais, univariados e multivariados, que serão aplicados à variável óbito por causas externas.

# Capítulo 1

## Inferência Bayesiana

Suponha que temos como objetivo obter informação sobre uma quantidade de interesse  $\theta$ , que não é uma quantidade diretamente observável mas está relacionada com uma amostra de dados. Antes de observar essa amostra, e dependendo do conhecimento de cada pesquisador, podemos ter densidades probabilísticas distintas para  $\theta$ , definidas por uma função de probabilidade  $p(\theta)$ , associadas a diferentes graus de incerteza. A idéia da inferência Bayesiana é reduzir a incerteza a respeito deste parâmetro através da amostra observada. Após observar um amostra  $x$  de um vetor aleatório  $X$  relacionado com  $\theta$ , podemos atualizar o conhecimento a respeito desse parâmetro através da seguinte expressão:

$$p(\theta|x) = \frac{p(\theta, x)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta, x)d\theta} \propto p(x|\theta)p(\theta). \quad (1.1)$$

Esse resultado é conhecido como Teorema de Bayes. Fica claro que através deste teorema podemos atualizar o conhecimento a respeito de  $\theta$ : dado o conhecimento prévio a respeito de  $\theta$  descrito probabilisticamente em  $p(\theta)$ , que denominamos distribuição a priori, combinamos com a verossimilhança  $p(x|\theta)$ , que nos indica o quão verossímil é uma amostra  $x$  dado um determinado valor de  $\theta$ , e assim o conhecimento a respeito desse parâmetro pode ser atualizado e descrito pela distribuição a posteriori  $p(\theta|x)$ .

## 1.1 Distribuição a Priori

Um dos temas mais polêmicos em Inferência Bayesiana é a especificação de uma distribuição a priori para o parâmetro de interesse  $\theta$ . Um conhecimento prévio a respeito deste parâmetro antes do experimento pode levar a diversas especificações de distribuições a priori, quando feitas por diferentes pesquisadores, conduta muito criticada por alguns estatísticos que ressaltam que esta subjetividade pode interferir imprudentemente nos resultados. Uma forma encontrada para contornar este problema é a especificação de prioris não informativas, que levam a resultados de inferência similares aos obtidos pela estatística clássica. Nessa seção descreveremos algumas das especificações a priori mais utilizadas: as prioris não informativas; as prioris conjugadas, que são muito utilizadas por levarem a cálculos simples da posteriori; e as prioris hierárquicas.

### 1.1.1 Prioris Conjugadas

Se  $F = \{p(x|\theta), \theta \in \Theta\}$  é uma classe de distribuições amostrais, dizemos que uma classe de distribuições  $P$  é conjugada a  $F$  se para todo  $p(x|\theta) \in F$  e  $p(\theta) \in P$  temos  $p(\theta|x) \in P$ , ou seja, a distribuição amostral e a priori pertencem a uma mesma família de distribuições. Nesse caso, obtemos a posteriori com uma simples atualização dos parâmetros da distribuição a priori, que são chamados de hiperparâmetros. É importante ressaltar que em geral não estamos interessados em fazer inferência sobre os hiperparâmetros.

### 1.1.2 Prioris não Informativas

Quando não existe um conhecimento prévio sobre o parâmetro de interesse  $\theta$ , é bastante comum atribuir uma distribuição a priori não informativa para esse parâmetro que não influencia no resultado final da inferência, levando em muitos casos a resultados similares aos obtidos na inferência clássica. Trabalhando com esta classe de prioris, o efeito da subjetividade por parte de cada pesquisador é eliminado, mas permanecemos com a vantagem da abordagem Bayesiana para fazer inferência.

Uma classe de prioris não informativas bastante utilizada foi proposta por Jeffreys

(1961). Para o cálculo da priori de Jeffreys, necessitamos primeiramente calcular a Informação de Fisher, que mede a curvatura média da função de verossimilhança. Quanto maior a Informação de Fisher, ou seja, maior a curvatura, maior é a informação contida na verossimilhança. Quando  $\theta$  é um escalar, temos a seguinte definição para informação de Fisher:

$$I(\theta) = E\left[-\frac{\partial^2 \log p(X|\theta)}{\partial \theta^2}\right],$$

onde  $E[X]$  denota a esperança da variável aleatória  $X$ . Caso  $\theta$  seja um vetor paramétrico, teremos uma matriz de informação de Fisher de  $\theta$ , definida por:

$$I(\theta) = E\left[-\frac{\partial^2 \log p(X|\theta)}{\partial \theta \partial \theta'}\right].$$

A partir da definição da informação de Fisher, pode-se definir a priori não informativa de Jeffreys que tem função densidade definida por

$$p(\theta) \propto I(\theta)^{\frac{1}{2}},$$

para  $\theta$  escalar, ou, caso  $\theta$  seja um vetor paramétrico,

$$p(\theta) \propto \det|I(\theta)|^{\frac{1}{2}},$$

### 1.1.3 Prioris Hierárquicas

A idéia aqui é dividir a especificação da distribuição a priori em estágios. Além de facilitar a especificação, esta abordagem é natural em determinadas situações experimentais. A distribuição a priori de  $\theta$  depende dos valores dos hiperparâmetros  $\phi$ , e podemos especificar  $p(\theta|\phi)$  ao invés de  $p(\theta)$ . Além disso, ao invés de fixar valores para os hiperparâmetros podemos especificar uma distribuição a priori  $p(\phi)$  para eles, completando assim a especificação do segundo estágio da hierarquia. A distribuição a priori marginal de  $\theta$  pode ser então obtida por integração:

$$p(\theta) = \int p(\theta, \phi) d\phi = \int p(\theta|\phi)p(\phi) d\phi$$



## 1.2 Estimação

A melhor forma de descrever o comportamento do parâmetro  $\theta$  é através de sua distribuição a posteriori, por exemplo através de um gráfico da sua função densidade. Muitas vezes, entretanto, existe o interesse de se obter uma informação resumida a respeito do parâmetro. A partir da posteriori, podem ser calculadas estatísticas que resumem de maneira eficiente o parâmetro  $\theta$ , como média, moda e mediana. Estatísticas como essas são chamadas de estimadores pontuais. Além da estimação pontual, podemos também fazer estimação por intervalo, que tem a vantagem de nos mostrar a precisão associada à estimativa. Nessa subseção introduzimos os conceitos das estimações pontual e por intervalo.

### 1.2.1 Estimação Pontual

A escolha de um estimativa para um parâmetro pode ser encarado como um problema de decisão, onde é necessário avaliar os custos e benefícios de cada escolha. Migon e Gamerman (1999) especificam um problema de decisão através da descrição de três espaços:

1. Espaço de parâmetros  $\Theta$ ;
2. Espaço de resultados possíveis de um experimento  $\Omega$ ;
3. Espaço de possíveis ações  $A$ .

Uma regra de decisão  $\delta$  é uma função definida em  $\Omega$ , que assume valores em  $A$ , isto é,  $\delta : \Omega \rightarrow A$ . A cada decisão  $\delta$ , cada valor do possível parâmetro  $\theta$ , podemos associar uma perda  $L(\delta, \theta)$ , assumindo valores positivos. A cada função perda selecionada é associada uma medida de risco a decisão  $\delta$ , que é dada por  $R(\delta) = E_{\theta|x}[L(\delta, \theta)]$ , ou seja, é a perda esperada a posteriori também chamada de Estimador de Bayes.

As funções perda mais utilizadas são:

1. perda quadrática:  $L(\delta, \theta) = (\delta - \theta)^2$ ;

2. perda absoluta:  $L(\delta, \theta) = |\delta - \theta|$ ;

3. perda zero e um:  $L(\delta, \theta) = \lim_{\epsilon \rightarrow 0} I_{|\theta - \delta|}([\epsilon, \infty])$ , onde  $I_x(A)$  é a função que assume o valor 1 se  $x \in A$  e 0 caso contrário.

A função perda pode ser interpretada como uma penalização pela tomada de uma decisão  $\delta$  quando assumimos que  $\theta$  é o verdadeiro valor do parâmetro. Através do Estimador de Bayes é selecionada a estimativa que minimiza a perda esperada. Pode-se mostrar que os Estimadores de Bayes associados às funções de perda quadrática, absoluta e zero e um, são a média, mediana e moda a posteriori, respectivamente.

### 1.2.2 Estimação por Intervalos

Um aspecto negativo da estimativa pontual é que ela resume a informação da distribuição a posteriori em um único ponto, não sendo possível portanto acessar a precisão deste resultado. Uma forma de contornar esse problema é através de intervalos de credibilidade para  $\theta$ .

Estamos interessados em saber com que probabilidade um parâmetro  $\theta$  desconhecido, que pode assumir valores em  $\Theta$ , pertence a um intervalo  $C$  (que também assume valores em  $\Theta$ ). Podemos formalizar o conceito de intervalo de credibilidade da seguinte forma:  $C$  é um intervalo de credibilidade de  $100(1 - \alpha)\%$ , ou nível de credibilidade (ou confiança)  $1 - \alpha$  para  $\theta$ , se  $P(\theta \in C) \geq 1 - \alpha$ .

Podemos construir inúmeros intervalos com a definição acima, mas em geral o interesse é encontrar um intervalo com o menor comprimento possível, que englobe os valores mais plausíveis para  $\theta$ , ou seja, aqueles com maior densidade a posteriori. Esses intervalos são chamados de intervalos de máxima densidade à posteriori, e podem ser especificados através da seguinte definição: Um intervalo de credibilidade  $C$  de  $100(1 - \alpha)\%$  para  $\theta$  é de máxima densidade a posteriori se  $C = \{\theta \in \Theta : p(\theta|x) \geq k(\alpha)\}$  onde  $k(\alpha)$  é a maior constante tal que  $P(\theta \in C) \geq 1 - \alpha$ .

No próximo capítulo será feita uma revisão de modelos lineares, onde será feita uma breve revisão teórica sobre modelos lineares Gaussianos, modelos dinâmicos lineares, modelos lineares generalizados e modelos lineares generalizados dinâmicos.

# Capítulo 2

## Modelos Lineares

### 2.1 Modelos Lineares Gaussianos

O apelo dos modelos de regressão é explicar uma ou várias variáveis de interesse (objetivo) em função de outras variáveis (explicativas). Em alguns casos, parte da variabilidade da variável de interesse  $y$ , onde  $y' = (y_1, \dots, y_n)$  pode ser explicada por um conjunto de variáveis explicativas, através de uma função linear. Seja  $X_1, \dots, X_p$  um conjunto de variáveis explicativas,  $p < n$ , podemos estabelecer a seguinte relação com  $y$ :

$$\begin{aligned}y_i &= \mu_i + e_i \\ \mu_i &= \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi},\end{aligned}$$

para  $i = 1, \dots, n$ , onde assumimos que os erros  $e_i$  são independentes e tem distribuição Normal com média zero e variância  $\sigma^2$ , e  $\mu_i$  é a parte determinística do modelo. Note que o valor esperado de  $y$  é a parte determinística do modelo, ou seja,  $E(y_i) = \mu_i$ ,  $i = 1, \dots, n$ .

Definindo  $\beta = (\beta_1, \dots, \beta_p)$ , podemos escrever a função de verossimilhança dos parâmetros  $\sigma^2$  e  $\beta$  como:

$$l(\beta, \sigma^2 | y) \propto \sigma^{-n} \exp \left\{ -\frac{(X\beta)'(X\beta)}{2\sigma^2} \right\}.$$

Fazendo modelagem sob o enfoque Bayesiano, é necessário também especificar prioris para os parâmetros  $\sigma^2$  e  $\beta$  para total especificação do modelo. Usualmente são atribuídas

prioris conjugadas normais para cada  $\beta_i$ ,  $i = 1, \dots, p$ , e Gama invertida para  $\sigma^2$ . Essas especificações deixam a priori e a distribuição amostral conjugadas, facilitando a inferência dos parâmetros.

## 2.2 Modelos Dinâmicos Lineares

Os modelos dinâmicos lineares (MDL) foram introduzidos por Harrison e Stevens (1976) e estão bem documentados em West e Harrison (1997). A idéia aqui é introduzir uma fonte extra de variabilidade no modelo, permitindo que coeficientes de regressão variem de forma suave ao longo do tempo. O modelo pode ser caracterizado por um par de equações: a equação observacional e a equação do sistema, que são dadas respectivamente por:

$$y_t = X_t' \beta_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_t^2) \quad (2.1)$$

$$\beta_t = G_t \beta_{t-1} + w_t, \quad w_t \sim N(0, W_t). \quad (2.2)$$

para  $t = 1, \dots, T$ , onde  $y_t$  é uma seqüência de observações realizadas no tempo  $t$ , condicionalmente independente à seqüência de parâmetros  $\beta_t$ ;  $X_t$  é um vetor de variáveis explicativas conhecidas no tempo  $t$ ,  $\beta_t$  é um vetor de dimensão  $p \times 1$ ,  $G_t$  é uma matriz de evolução dos parâmetros e tem dimensão  $p \times p$ ,  $\sigma_t^2$  e  $W_t$  são as variâncias de  $\epsilon_t$  e  $w_t$ , respectivamente, sendo  $\sigma_t^2$  um escalar e  $W_t$  uma matriz  $p \times p$ . Um modelo dinâmico linear é completamente especificado pela quádrupla  $\{X_t, G_t, \sigma_t^2, W_t\}$ . A escolha de  $X_t$  e  $G_t$  depende do modelo e da natureza dos dados que estão sendo analisados. Casos particulares dos modelos dinâmicos lineares incluem o modelo de regressão (fazendo  $G_t = I_p$ , a matriz identidade de ordem  $p$ , e  $W_t = 0$ ), e modelos lineares de séries temporais (fazendo  $X_t = X, G_t = G, \sigma_t^2 = \sigma^2$  e  $W_t = W$ ), tais como os processos ARIMA de Box e Jenkins (1976).

### Inferência sob modelos dinâmicos lineares

Os aspectos de inferência nos modelos dinâmicos lineares seguem os passos usuais da estatística Bayesiana, explorando seu aspecto sequencial e combinando duas operações

principais a “evolução” para construir a priori, e a “atualização” para incorporar uma nova observação no tempo  $t$ . Vamos denotar toda informação obtida até o período de tempo  $t$  por  $D_t = D_{t-1} \cap y_t$ .  $D_0$  portanto representa a informação a priori. Para cada período de tempo  $t$ , as distribuições a priori, preditiva e posteriori são dadas respectivamente por:

$$\begin{aligned} p(\beta_t|D_{t-1}) &= \int p(\beta_t|\beta_{t-1})p(\beta_{t-1}|D_{t-1})d\beta_{t-1} \\ p(y_t|D_{t-1}) &= \int p(y_t|\beta_t)p(\beta_t|D_{t-1})d\beta_t \\ p(\beta_t|D_t) &\propto p(\beta_{t-1}|D_{t-1})p(y_t|D_{t-1}) \end{aligned}$$

sendo a última equação obtida através do teorema de Bayes em (1.1). Quando  $\{\mathbf{X}_t, \mathbf{G}_t, \sigma_t^2, \mathbf{W}_t\}$  é conhecido a constante de integração na especificação acima pode ser facilmente obtida através do Filtro de Kalman (Anderson e Moore, 1979). Neste caso, dados os valores  $y_1, \dots, y_{t-1}$ , é fácil prever  $y_t$  e estimar os parâmetros de estado não observáveis  $\beta_1, \dots, \beta_{t-1}$  através das distribuições preditiva ( $y_t|D_{t-1}$ ) e posteriori ( $\beta_t|D_{t-1}$ ), respectivamente. Quando  $y_t$  é observado, a estimativa de  $\beta_{t-1}$  pode ser atualizada usando sua distribuição a posteriori dado  $D_t$ . Considere o modelo dinâmico linear com  $\sigma_t^2$  desconhecido. Por simplicidade fazemos  $\sigma_t^2 = \sigma^2$ , para  $\forall t$ . Vamos denotar a distribuição a posteriori de  $\beta$  no tempo  $t - 1$  por  $(\beta_{t-1}|D_{t-1}, \sigma^2) \sim N(m_{t-1}, \sigma^2 C_{t-1})$  e a distribuição a posteriori de  $\phi = \sigma^{-2}$  por  $\phi|D_{t-1} \sim G(n_{t-1}/2, n_{t-1}s_{t-1}/2)$ . Além disso, supomos que a priori inicial em  $t = 0$  é normal:  $(\beta_0|D_0) \sim N(m_0, C_0)$ , para algum vetor de médias  $m_0$  e matriz de covariâncias  $C_0$  conhecidos. Temos então que:

1. Condicionalmente a  $\sigma^2$ :

- **Evolução:** Temos que no tempo  $t$  a distribuição a priori para  $\beta_t$  é

$$\beta_t|\sigma^2, D_{t-1} \sim N(a_t, \sigma^2 R_t),$$

onde,  $a_t = G_t m_{t-1}$  e  $R_t = G_t C_{t-1} G_t' + W_t$ .

- **Previsão:** Temos que a distribuição preditiva um passo a frente é

$$y_t|\sigma^2, D_{t-1} \sim N(f_t, \sigma^2 Q_t),$$

onde,  $f_t = X_t' a_t$  e  $Q_t = X_t' R_t X_t + 1$ .

- **Atualização:** A distribuição a posteriori no tempo  $t$  é dada por

$$\beta_t | D_t \sim N(m_t, \sigma^2 C_t),$$

com  $m_t = a_t + A_t \epsilon_t$  e  $C_t = R_t - A_t A_t' Q_t$ , onde  $A_t = R_t X_t' / Q_t$   
e  $\epsilon_t = y_t - f_t$

A precisão  $\phi$  é atualizada pela seguinte relação

$$\phi | D_t \sim G(n_t/2, n_t s_t/2),$$

onde  $n_t = n_{t-1} + 1$  e  $n_t s_t = n_{t-1} s_{t-1} + \epsilon_t^2 / Q_t$ .

2. Incondicionalmente a  $\sigma^2$ :

- **Evolução:** Temos que no tempo  $t$  a distribuição a priori para  $\beta_t$  é

$$\beta_t | D_{t-1} \sim t_{n-1}(a_t, s_{t-1} R_t),$$

onde  $a_t$  e  $R_t$  são definidos como anteriormente, e  $s_{t-1}$  é a estimativa da variância no tempo  $t - 1$ .

- **Previsão:** Temos que a distribuição preditiva um passo a frente é

$$y_t | D_{t-1} \sim t_{n-1}(f_t, Q_t^*),$$

onde  $f_t$ ,  $Q_t$  e  $s_{t-1}$  são definidos como anteriormente e  $Q_t^* = s_{t-1} Q_t$ .

- **Atualização:** A distribuição a posteriori no tempo  $t$  é dada por

$$\beta_t | D_t \sim t_{n-1}(m_t, s_t C_t),$$

onde  $m_t$  e  $C_t$  são definidos como anteriormente e  $s_t$  é a estimativa da variância no tempo  $t$ .

## 2.3 Modelos Lineares Generalizados

Algumas vezes, a hipótese de normalidade da variável de interesse, que é assumida nos modelos lineares gaussianos, não é razoável. Por exemplo podemos citar variáveis resposta positivas discretas, que são observadas em forma de contagens, como o número de casos reportados de uma certa doença ou número de casos de morte por violência. Outros exemplos são variáveis que assumem resposta do tipo zero ou um e aquelas que assumem valores no intervalo entre zero e um.

Uma possível alternativa é utilizar os modelos lineares generalizados, que foram introduzidos por Nelder e Wedderburn (1972) (veja também McCullagh e Nelder, 1989) que são uma generalização dos modelos lineares gaussianos apresentados anteriormente. Assumimos agora que as observações vem de uma distribuição da família exponencial. Para utilização destes modelos é necessário assumir que a variável de interesse tenha uma distribuição que pertença a família exponencial. Na notação utilizada por McCullagh e Nelder (1989), dizemos que  $y$  pertence à família exponencial se a sua densidade pode ser escrita da forma abaixo:

$$f(y; \theta, \phi) = \exp \left\{ \frac{1}{a(\phi)} [y\theta - b(\theta)] + c(y; \phi) \right\},$$

onde  $\theta$  é o parâmetro natural ou canônico,  $\phi$  é um parâmetro de escala positivo, e  $b(\cdot)$  e  $c(\cdot)$  são funções conhecidas. Pode ser mostrado que a média e a variância de uma variável aleatória  $y$  cuja distribuição pertence à família exponencial, na forma canônica usada por McCullagh e Nelder (1989), são dadas por:

$$E(y) = b'(\theta) = \mu \tag{2.3}$$

$$Var(y) = a(\phi)b''(\theta) \tag{2.4}$$

A idéia dos modelos lineares generalizados é estabelecer uma relação entre um componente sistemático  $\eta = X\beta$  e o componente aleatório  $\mu$  (média da variável resposta) através de uma função de ligação monótona derivável  $g(\cdot)$ , tal que  $\eta = g(\mu)$ .

Para ilustrar a teoria descrita nesta seção, a seguir será dado um exemplo com a distribuição Poisson. Podemos escrever a função de probabilidade de uma variável aleatória  $y$  com distribuição de Poisson com média  $\lambda$  da seguinte forma:

$$f(y; \lambda, \phi) = \frac{e^{-\lambda} \lambda^y}{y!} = \exp \{y \log \lambda - \lambda - \log(y!)\}$$

Dessa forma, verificamos ser possível escrever a verossimilhança de  $y$  no formato da família exponencial, mostrando que a distribuição de Poisson pertence à essa família. Comparando a expressão acima com a notação geral da família exponencial de McCullagh e Nelder (1989), temos que  $a(\phi) = 1$ ,  $\theta = \log(\lambda)$ ,  $b(\theta) = \lambda$  e  $c(y; \phi) = 1$ .

Note que utilizando as equações 2.3 e 2.4, podemos verificar facilmente que a média e variância de  $y$  são iguais a  $\lambda$ , notando que a função de ligação  $g(\cdot)$  utilizada foi a logarítmica.

## 2.4 Modelos Lineares Dinâmicos Generalizados

Baseado nos modelos lineares generalizados de Nelder e Wedderburn (1972), os modelos lineares generalizados dinâmicos foram propostos por West, Harrison e Mignon (1985), permitindo observações na família exponencial como uma extensão aos Modelos Dinâmicos Lineares. A equação de observação em (2.1) é substituída por

$$f(y_t; \theta_t, \phi_t) = \exp \left\{ \frac{1}{a(\phi_t)} [y_t \theta_t - b(\theta_t)] + c(y_t; \phi_t) \right\},$$

e, seguindo a mesma idéia dos modelos lineares generalizados, é estabelecida uma função de ligação que relaciona a média  $\mu_t = E[y_t | \theta_t] = b'(\theta_t)$  e as covariáveis  $X_t$ , através de  $g(\mu_t) = \theta_t = X_t' \beta_t$ . A equação do sistema (em 2.2) permanece a mesma:

$$\beta_t = G_t \beta_{t-1} + w_t, \quad w_t \sim N(0, W_t).$$

Para fazer inferência a respeito de  $\beta_t$ , é necessário, primeiramente, especificar uma distribuição a priori para  $\beta_0$ , que pode ser por exemplo uma priori conjugada. Para obter a distribuição a posteriori dos parâmetros de estado, West, Harrison e Mignon (1985) sugerem o uso do estimador linear de Bayes. Outra forma de estimar é através dos métodos



computacionalmente intensivos MCMC (Markov Chain Monte Carlo), que serão apresentados no capítulo 4 dessa dissertação.

No próximo capítulo serão abordadas alternativas para modelar o componente estocástico, quando a suposição de independência espacial não é razoável, para modelar uma possível dependência espacial serão propostos Campos Aleatórios Markovianos Gaussianos univariados e multivariados.

# Capítulo 3

## Modelos estruturados no espaço

Seja  $y(s)$  uma variável aleatória em  $s \in D \subset R^2$ . Suponha que uma amostra dessa variável é obtida pela medição de  $y$  em  $N$  locais amostrais  $\{s_1, s_2, \dots, s_N\}$ . Podemos escrever  $y(s_i)$ ,  $i, \dots, N$  como  $y(s_i) = \mu(s_i) + \phi(s_i)$ , onde  $\mu = (\mu(s_1), \dots, \mu(s_N))$  são médias que podem ou não depender dos locais de observação  $s_i, i = 1, \dots, N$  e  $\phi = (\phi(s_1), \phi(s_2), \dots, \phi(s_N))$ , são erros (componente estocástico do modelo).

Espera-se que medições feitas em regiões próximas sejam similares. É importante ressaltar porém que a noção de proximidade é muito subjetiva, dependendo do problema abordado. Em determinados casos,  $\mu$  pode ser explicada por uma ou mais covariáveis como em um problema de regressão linear simples ou múltiplo, onde assumimos os erros  $\phi(s_i)$ 's como sendo independentes. Caso o componente estocástico apresente uma estrutura espacial, entretanto, não podemos assumir a independência desses erros. Uma alternativa para este tipo de problema é utilizar os Campos Aleatórios Markovianos Gaussianos, que serão apresentados a seguir.

### 3.1 Campos Aleatórios Markovianos Gaussianos

Caso o componente estocástico  $\phi = (\phi_1, \dots, \phi_N)$  não satisfaça a condição de independência e exista correlação espacial na área de estudo, é necessário especificar uma estrutura que acomode esta dependência espacial. Uma especificação muitas vezes utilizada para  $\phi$  é a de um Campo Aleatório Markoviano Gaussiano (CAMG), que tem a

seguinte notação:

$$\phi \sim CAMG(\mu, P),$$

onde  $\mu$  é um vetor de médias  $\mu = (\mu_1, \dots, \mu_N)$ , e  $P$  é uma matriz de precisões, tal que

$$p(\phi) \propto \exp \left\{ -\frac{1}{2}(\phi - \mu)'P(\phi - \mu) \right\}$$

Um caso particular desses modelos são os modelos propostos por Besag et.al. (1991), que podem ser obtidos fazendo  $P = M$ , onde:

$$M_{ij} = \begin{cases} m_i & , \quad i = j \\ -w_{ij} & , \quad i \sim j, \\ 0 & , \quad \text{caso contrário} \end{cases}$$

onde para  $i, j = 1, \dots, N$ ,  $i \sim j$  denota que as regiões  $s_i$  e  $s_j$  são vizinhas,  $m_i$  é o número de vizinhos da região  $i$  e  $w_{ij} > 0$  é uma medida (ou peso) de similaridade entre as regiões  $s_i$  e  $s_j$ . Os critérios para a escolha dos pesos  $w_{ij}$  podem ser diversos. Um dos critérios mais utilizados é baseado em fronteiras, dando peso  $w_{ij} = 1$  se o município  $i$  faz fronteira com o município  $j$ , e  $w_{ij} = 0$  caso contrário. Outro critério bastante utilizado é igualar  $w_{ij}$  ao inverso da distância entre os centróides dos município  $i$  e  $j$ .

A matriz  $M$  especificada é singular, portanto sua inversa a matriz de covariâncias, não existe. Como consequência, a distribuição conjunta de  $\phi$  não é própria quando  $P = M$ . Para contornar esse problema, Vivar-Rojas (2004) propõe uma modificação dos modelos de Besag et. al. (1991), definindo uma nova matriz de precisões  $P$ . A matriz de precisões modificada é definida por  $P = \lambda(\alpha I_N + M)$ , onde  $\lambda$  é um parâmetro de escala,  $I_N$  é uma matriz identidade (N x N) e  $\alpha$  é o parâmetro que controla o grau de correlação espacial corrigindo a singularidade da matriz  $M$ . Com a adição de  $\alpha$  à diagonal de  $M$ , a matriz  $P$  passa ser dominada pela diagonal e, por consequência, passa a ser positiva definida (Harville, 1997) tornando a distribuição de  $\phi$  própria. Uma forma ainda mais geral para a especificação de uma distribuição própria para  $\phi$  pode ser vista em Gamerman, Moreira e Rue (2003).

No modelo especificado por Vivar-Rojas (2004), temos:

$$\phi_i | \phi_j, \sigma^2 \sim N \left( \frac{\alpha + \mu_i \sum_{j:j \sim i} w_{ij} \phi_j}{\alpha + m_i}, \frac{\sigma^2}{(\alpha + m_i)} \right), \quad j \neq i$$

No caso particular do modelo de Besag, et. al. (1991),  $\alpha = 0$  e denotamos a distribuição de  $\phi$  por Modelo Condicional Autoregressivo Gaussiano Intrínseco (CAR). Esta distribuição condicional leva à seguinte distribuição conjunta para  $\phi$ :

$$p(\phi | \sigma^2) \propto \frac{1}{\sigma^n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{i < j} w_{ij} (\phi_i - \phi_j)^2 \right\}.$$

Besag e Kooperberg (1995) demonstraram que impondo que a soma dos efeitos aleatórios seja igual a zero, e especificando um intercepto com locação invariante e priori Uniforme  $(-\infty, \infty)$ , que é equivalente a uma nova parametrização do modelo incluindo um intercepto, garante-se a identificabilidade do modelo.

## 3.2 Modelos *CAR* Multivariados

Os *CAR* multivariados são uma extensão multivariada do modelo *CAR* apresentado anteriormente. Como nos modelos *CAR*, a sua versão multivariada pode ser utilizada para modelar efeitos aleatórios, e é uma distribuição imprópria. Por simplicidade, vamos supor que desejamos modelar duas variáveis aleatórias.

Defina  $w_{i+} = \sum_j w_{ij}$ , ou seja,  $w_{i+}$  é a soma dos pesos da matriz  $W$  na linha  $i$ , e seja  $D$  uma matriz diagonal  $N \times N$  ( $N$  é o número de áreas geográficas) cuja diagonal é formada pelos elemento  $w_{i+}$ . Denotamos por  $D = \text{Diag}(w_{i+})$ . Os efeitos aleatórios representados pelas duas variáveis têm a seguinte representação:  $\Phi' = (\phi'_1, \phi'_2)$ , sendo  $\phi'_1 = (\phi_{11}, \dots, \phi_{1N})$  e  $\phi'_2 = (\phi_{21}, \dots, \phi_{2N})$ . A distribuição conjunta multivariada do *CAR* é definida como:

$$p(\Phi) \propto \exp \left\{ -\frac{1}{2} \Phi' [\Lambda \otimes (D - W)] \Phi \right\}, \quad (3.1)$$

onde  $\Lambda$  é uma matriz  $2 \times 2$  positiva definida e  $\otimes$  é o produto de Kronecker.

A distribuição condicional para  $(\phi_{i1}, \phi_{i2})$  é

$$(\phi_{i1}, \phi_{i2}) | \phi_{-(i1, i2)} \sim N((\bar{\phi}_{i1}, \bar{\phi}_{i2}), (w_{i+} \Lambda)^{-1}) \quad (3.2)$$

onde  $\phi_{-(i1, i2)}$  é a sequência de todos os  $\phi'_{ij}$ s exceto  $\phi_{i1}$  e  $\phi_{i2}$ , e  $\bar{\phi}_{i1}$  e  $\bar{\phi}_{i2}$  são as médias dos efeitos aleatórios para cada área especificada pelos vizinhos para as variáveis 1 e 2 respectivamente, ou seja,  $\bar{\phi}_{i1} = \sum_j w_{ij} \phi_{j1} / w_{i+}$  e  $\bar{\phi}_{i2} = \sum_j w_{ij} \phi_{j2} / w_{i+}$ .

A precisão da distribuição condicional em (3.2) é controlada pelo parâmetro de escala  $w_{i+}$ , onde áreas com mais vizinhos levam a uma maior precisão condicional para os  $\phi_{i1}$  e  $\phi_{i2}$  correspondentes. Definindo  $\Sigma = \Lambda^{-1}$ , a matriz de covariâncias da distribuição condicional é dada por  $\frac{1}{w_{i+}} \Sigma$ , onde a correlação condicional entre  $\phi_{i1}$  e  $\phi_{i2}$ ,  $i = 1, \dots, N$ , é dada por  $\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_{11}\sigma_{22}}}$ . No WinBugs, software utilizado para fazer a inferência dos parâmetros, está implementando o CAR univariado e multivariado.

Nessa dissertação os modelos propostos não tem uma distribuição a posteriori que pode ser obtida analiticamente, então serão utilizados métodos computacionalmente intensivos, que serão vistos no próximo capítulo.

# Capítulo 4

## Métodos Computacionalmente Intensivos

### 4.1 Introdução

Muitas vezes a distribuição posteriori dos parâmetros de um modelo não pode ser obtida analiticamente. Nesses casos, uma solução é obter amostras dessas distribuições através de métodos computacionais tais como o Método de Monte Carlo via Cadeias de Markov (MCMC) (ver Gamerman e Lopes, 2006). Nessa seção será feita uma revisão dos principais algoritmos de MCMC. Vamos definir  $\Theta = \{\theta_1, \dots, \theta_d\}$  como sendo a coleção de parâmetros desconhecidos do modelo.

### 4.2 Algoritmo Metropolis-Hastings

Os algoritmos de Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970) usam a mesma idéia dos métodos de rejeição, onde uma distribuição auxiliar é utilizada para a geração de uma cadeia de pontos, que são aceitos com uma determinada probabilidade. Este mecanismo de correção garante a convergência da cadeia para a distribuição de equilíbrio, que neste caso é a distribuição a posteriori de  $\Theta$ .

Suponha que a cadeia esteja no estado  $\Theta^j$  e um valor  $\Theta'$  é gerado de uma distribuição proposta  $q(\cdot|\Theta^j)$ . Note que a distribuição proposta pode depender do estado atual da

cadeia. Por exemplo,  $q(\cdot|\Theta^j)$  poderia ser uma distribuição Normal centrada em  $\Theta^j$ . O novo valor  $\Theta'$  é aceito com probabilidade dada abaixo:

$$\alpha(\Theta^j, \Theta') = \min \left\{ 1, \frac{\pi(\Theta')q(\Theta^j|\Theta')}{\pi(\Theta^j)q(\Theta'|\Theta^j)} \right\},$$

onde  $\pi(\Theta)$  denota a densidade a posteriori de  $\Theta$ .

Uma característica importante é que só precisamos conhecer  $\pi$  parcialmente, isto é, a menos de uma constante, já que neste caso a razão acima não se altera. Isto é fundamental em aplicações Bayesianas onde não conhecemos completamente a distribuição a posteriori. Note também que a cadeia pode permanecer no mesmo estado por muitas iterações. Na prática costuma-se monitorar a porcentagem das iterações para quais novos valores são aceitos, assim a função  $q(\cdot|\Theta^j)$  pode ser ajustada para que esta porcentagem não seja muito alta ou muito baixa.

Em termos práticos, o algoritmo de Metropolis-Hastings pode ser especificado pelos seguintes passos:

1. Inicialize o contador de iterações  $t = 0$  e especifique um valor inicial  $\Theta^0$ ;
2. Gere um novo valor  $\Theta'$  da distribuição  $q(\cdot|\Theta^t)$ ;
3. Calcule a probabilidade de aceitação  $\alpha(\Theta^t, \Theta')$  e gere  $u \sim U(0, 1)$ ;
4. Se  $u \leq \alpha(\Theta^t, \Theta')$  então aceite o novo valor e faça  $\Theta^{t+1} = \Theta'$ , caso contrário rejeite e faça  $\Theta^{t+1} = \Theta^t$ ;
5. Incremente o contador de  $t$  para  $t + 1$  e volte ao passo 2.

O algoritmo deve ser executado até a convergência da cadeia e a partir daí os pontos gerados podem ser considerados como uma amostra da distribuição a posteriori.

### 4.3 Metropolis-Hastings com uma variável por vez

Esse algoritmo é similar ao algoritmo de Metropolis-Hastings, porém cada parâmetro  $\theta_1, \dots, \theta_d$  é atualizado separadamente, sendo testados sucessivamente. O algoritmo é especificado pelos seguintes passos:

1. Faça  $t = 0$ ,  $i = 1$  e especifique um valor inicial  $\Theta^0$ ;
2. Gere um novo valor  $\theta'_i$  da distribuição  $q(\cdot|\theta_i^t, \Theta_{-i}^t)$ , definindo  $\Theta_{-i}^t = (\theta_1^{(t+1)}, \dots, \theta_{i-1}^{(t+1)}, \theta_{i+1}^{(t)}, \dots, \theta_d^{(t)})$ ;
3. Calcule a probabilidade de aceitação  $\alpha(\theta_i^t, \theta'_i, \Theta_{-i}^t)$  e gere  $u \sim U(0, 1)$ ;
4. Se  $u \leq \alpha(\theta_i^t, \theta'_i, \Theta_{-i}^t)$  então aceite o novo valor e faça  $\theta_i^{t+1} = \theta'_i$ , caso contrário rejeite e faça  $\theta_i^{t+1} = \theta_i^t$ ;
5. Se  $i < d$  faça  $i = i + 1$  e retorne ao passo 3;
6. Incremente o contador de  $t$  para  $t + 1$  e volte ao passo 2.

Note que cada iteração se completa após  $d$  movimentos ao longo dos eixos coordenados das componentes de  $\theta$ . Assim como no caso anterior, o algoritmo deve ser executado até a convergência da cadeia, e a partir daí os pontos gerados podem ser considerados como uma amostra da distribuição a posteriori.

## 4.4 Amostrador de Gibbs

No amostrador de Gibbs (Geman e Geman, 1984; Gelfand e Smith, 1990) a cadeia irá sempre se mover para um novo valor, ou seja, não existe mecanismo de aceitação e rejeição. As transições de um estado para o outro são feitas através da geração de pontos das distribuições condicionais completas  $\pi(\theta_i|\Theta_{-i})$ , onde  $\Theta_{-i} = \{\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n\}$ . Note que a distribuição condicional completa é a distribuição da  $i$ -ésima componente condicionada a todas as outras componentes. Ela pode ser obtida a partir da distribuição conjunta, fazendo

$$\pi(\theta_i|\Theta_{-i}) = \frac{\pi(\Theta)}{\int \pi(\Theta) d\theta_i}.$$

Em muitas situações a geração de uma amostra diretamente de  $\pi(\Theta)$  pode ser custosa, complicada ou simplesmente impossível. Entretanto, se as distribuições condicionais completas forem completamente conhecidas, então podemos utilizar o amostrador de Gibbs que é definido pelo seguinte esquema:



1. Faça  $t = 0$ ,  $i = 1$ , e especifique um valor inicial  $\Theta^0$ ;
2. Gere um novo valor  $\theta_i^{t+1}$  da distribuição  $q(\cdot|\Theta_{-i}^t)$ , que é a densidade da distribuição condicional completa de  $\theta_i$ ;
3. Se  $i < d$  faça  $i = i + 1$  e retorne ao passo 3;
4. Incremente o contador de  $t$  para  $t + 1$  e volte ao passo 2.

Após a convergência, os valores resultantes formam uma amostra de  $\pi(\Theta)$ . Vale notar que mesmo em problemas de grandes dimensões todas as simulações podem ser univariadas, o que as vezes pode ser uma vantagem computacional. Note também que o amostrador de Gibbs é um caso especial do algoritmo de Metropolis-Hastings com uma variável por vez tomando a distribuição condicional completa como proposta, que leva a uma probabilidade de aceitação igual a 1.

## 4.5 Diagnósticos de Convergência

Sabemos que a partir da convergência, os pontos gerados da cadeia podem ser vistos como pontos gerados da distribuição a posteriori da cadeia. Os pontos gerados antes da convergência devem ser descartados e os períodos iniciais onde foi feita a geração desses pontos descartados é denominado de período de aquecimento da cadeia ou *burn-in*. Uma questão importante nos algoritmos de MCMC é saber quantas iterações são necessárias para o período de aquecimento. Para isso utilizam-se diagnósticos de convergência, tais como o de Geweke (1992) e Gelman e Rubin (1992) que podem ser vistos em Gamerman e Lopes (2006). Muitas vezes é possível verificar a convergência através de uma análise inspeção visual, utilizando histogramas e densidades Kernel da estimativa da distribuição a posteriori do parâmetro de interesse.

Outra questão importante é a verificação de uma possível autocorrelação das cadeia de pontos obtidos a partir do "burn-in". Cadeias autocorrelacionadas podem levar a uma subestimação da variância do parâmetro. Para corrigir esse problema podemos selecionar um ponto a cada  $k$  iterações (onde  $k$  depende do lag da autocorrelação) para formar a

amostra da posteriori do parâmetro em questão.

Para um melhor compreensão das variáveis óbitos por causas externas, valor esperado e a covariável escola, serão realizadas análises descritivas espaciais e temporais no estado do Paraná, no capítulo seguinte.

## Capítulo 5

# Análise do Número de Óbitos por Causas Externas no estado do Paraná

Atualmente a Segurança Pública do Estado do Paraná (SESP) passa por grandes questionamentos em relação ao seu desempenho no controle da violência. A opinião pública cobra um maior investimento do Estado, como o aumento efetivo de policiamento. Em contrapartida, o governo do Estado do Paraná defende-se com o argumento de que o Paraná tem um dos melhores quadros nacionais em relação a violência. A Secretária de Segurança Pública demonstra grande confiança em seu trabalho e alega que o Paraná é um dos únicos Estados a fornecer informação correta sobre o número de homicídios e outras mortes violentas, enquanto que em outros Estados, em geral, existe um alto índice de subnotificação, ou seja, divulgam números menores que a realidade.

No Estado do Paraná, a informação do número de homicídios e mortes violentas é feita através de um boletim unificado (BOU). A Polícia Civil e Militar trabalham de forma integrada no fornecimento de dados para o BOU, esta união ajuda com um fornecimento mais acurado de dados a respeito da situação de violência do Estado do Paraná. Esta discussão levou o Paraná a não divulgar os números de violência para a Secretarias Nacional de Segurança Pública (SENASP).

Utilizamos nesse trabalho dados fornecidos pelo Banco de dados do Sistema Único de

Saúde (DataSus), onde obtivemos a variável óbitos por causas externas. O DataSus faz o registro de casos no hospital, toda morte cuja causa não seja natural é classificada como óbito por causa externa, a vantagem de utilizar a informação do DataSus, é minimizar o problema de subnotificação da polícia, onde ocorre um número de registros menor que o real, Kahn (2007). Dessa forma, o número de óbitos por causas externas agrupa contagens do número de homicídios, suicídios e mortes por acidentes por município.

Ao longo dessa seção, a variável número de óbitos por causa externas será analisada de 1979 a 2004 no Estado do Paraná, com o objetivo de detectar padrões espaciais ao longo do tempo e encontrar regiões críticas que necessitem de uma atenção especial da Segurança Pública. Além disso, através desta evolução ao longo tempo é possível constatar se as medidas tomadas pelo governo tiveram um efeito desejado. Espera-se que regiões com um risco relativo alto em determinado ano tenham seu risco diminuído após uma ação da Segurança Pública. Para tentar explicar parte da variação do número de óbitos, é proposta a utilização da covariável logaritmo do número de escolas por município dividido por sua população, pois tirando o efeito da densidade populacional, espera-se que em locais com maior concentração de escolas menos óbitos por causas externas ocorram. Sendo assim, essa covariável está representando a educação como um fator de prevenção social. Analisamos também a variável número de casos de tráfico de drogas, que tem uma grande associação com óbitos por causas externas, principalmente com homicídio (ver Assunção, Beato, Silva e Marinho, 2001). Através de denúncias da população feita pelo 181, um serviço telefônico disponibilizado pela polícia, o cidadão consegue apontar os principais focos de tráfico de drogas de sua cidade. A variável utilizada nesse trabalho é o número de casos de tráfico de droga confirmados pela polícia, ou seja, nem toda denúncia da população torna-se um inquérito policial.

## 5.1 Aspectos Gerais

O Paraná é uma das 27 unidades federativas do Brasil. Está situado na região Sul do país e tem como limites São Paulo (a norte e nordeste), oceano Atlântico (leste), Santa Catarina (sul), Argentina (sudoeste), Paraguai (oeste) e Mato Grosso do Sul (noroeste).

Ocupa uma área de 199.314  $km^2$ . Sua capital é Curitiba e outras importantes cidades são Londrina, Maringá, Ponta Grossa, Paranavaí, Foz do Iguaçu, Cascavel, Guarapuava e Paranaguá.

A Figura 5.1 mostra o mapa do Paraná com 399 municípios destacando as cidades de maior porte populacional e econômico.

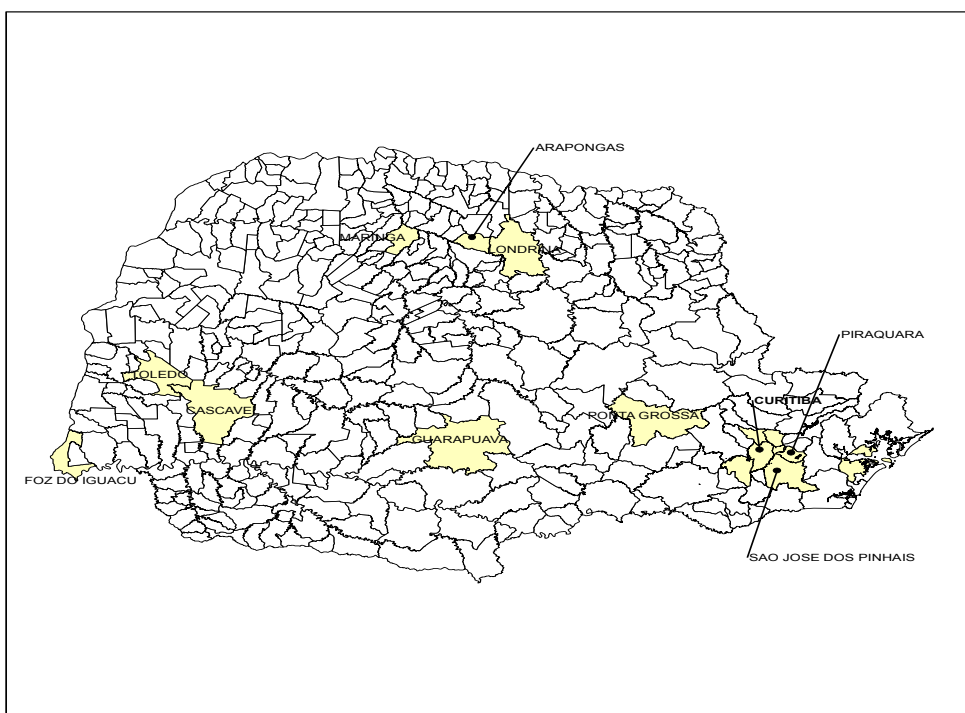


Figura 5.1: *Mapa do Estado do Paraná*

A Tabela 5.1 mostra os principais municípios analisados no mapa do Paraná com suas respectivas populações. No ano de 2004 essas cidades concentraram cerca de 50% de registros de óbitos por causas externas mostrando que o fator populacional, como é de se esperar, influencia de maneira relevante a distribuição de óbitos no estado do Paraná.

Cidade	População
Curitiba	1788559
Londrina	495696
Maringá	324397
Ponta Grossa	304973
São José dos Pinhais	261125

Tabela 5.1: População das principais cidades do Paraná

A Figura 5.2 mostra as variáveis número médio de óbitos por causas externas (considerando todo o Estado do Paraná), para os anos de 1980, 1990, 1996 e 2004. Pode-se observar um crescimento do número médio de óbitos por causas externas ao longo dos anos, que se acentuou nos últimos quatro anos analisados.

O aumento do número de óbitos devido a acidentes de trânsito deve-se ao aumento da frota de carros, sem o planejamento adequado por parte dos municípios para suportar esta demanda. O número de suicídios tem como origem razões emocionais e psicológicas, sendo difícil estabelecer métodos de prevenção. O número de homicídios tem diversas explicações sociais e econômicas, que devem ser levadas em consideração para entender a sua dinâmica. Dentre muitas explicações para o grande aumento de registros de mortes violentas, vale ressaltar a correlação que existe com o tráfico de drogas.

Na Figura 5.3 o número de óbitos por causas externas ao longo do tempo é apresentado individualmente para os municípios de Curitiba, Foz do Iguaçu e Londrina, locais com maior volume de óbito no Estado do Paraná. Podemos observar uma tendência suave de crescimento no período observado, de 1979 a 2004.

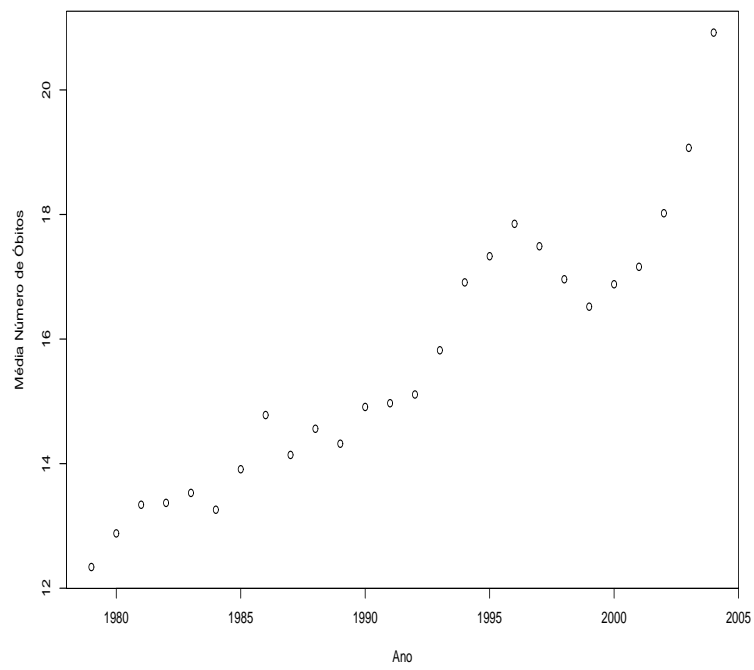


Figura 5.2: *Número médio de óbitos por causas externas versus Ano*

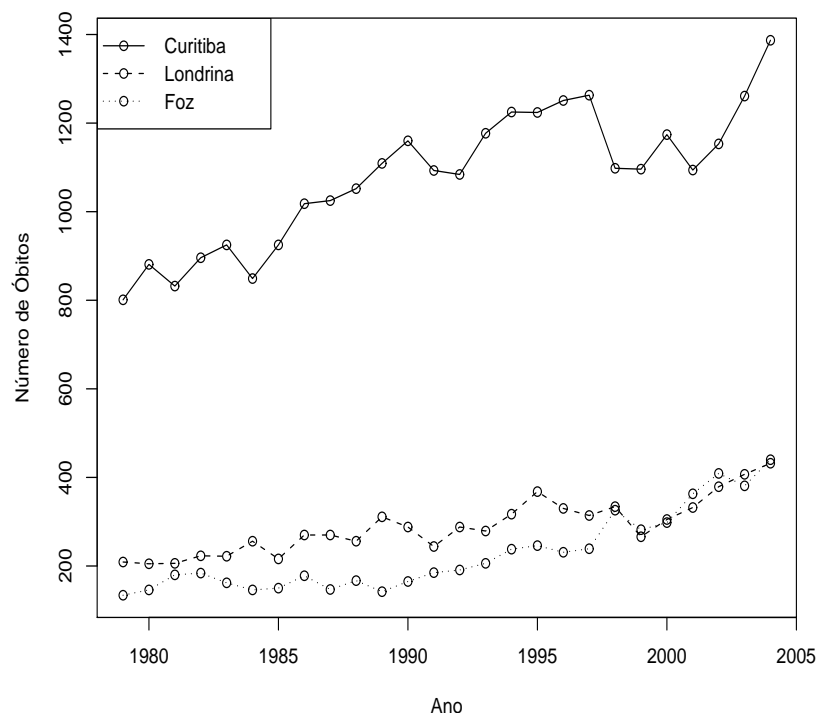


Figura 5.3: *Número de óbitos por causas externas das principais cidades do Estado do Paraná*

Além de analisar a evolução temporal do número de óbitos por causas externas, é relevante também analisar o comportamento dessa variável no espaço, para detectar possíveis padrões espaciais. Na Figura 5.4 temos a variável número médio de óbitos por causas externas, por município (média calculada considerando os 26 anos estudados), mostrado no mapa do estado do Paraná. Apesar de aparentemente existir bastante variabilidade dessa variável no espaço, é possível constatar um agrupamento espacial em torno da Região Metropolitana de Curitiba.

Como uma possível variável explicativa, é proposto o logaritmo do número de escolas em cada município dividido pela sua respectiva população. Na Figura 5.5 apresentamos a média do logaritmo do número de escolas em cada município (nos 26 anos estudados). Comparando visualmente as figuras 5.4 e 5.5, verificamos uma correlação positiva entre o número de óbitos e o número médio de escolas. Esse resultado não é inesperado, pois a



densidade populacional está fortemente correlacionada com ambas as variáveis. Tirando o efeito do tamanho populacional, espera-se que onde exista uma maior concentração de escolas, menos óbitos por violência ocorram.

Nas próximas análises, trabalharemos com o número de escolas dividido pela População do município, e dessa maneira poderemos analisar se o número de escolas por habitantes tem um impacto positivo no controle da violência.

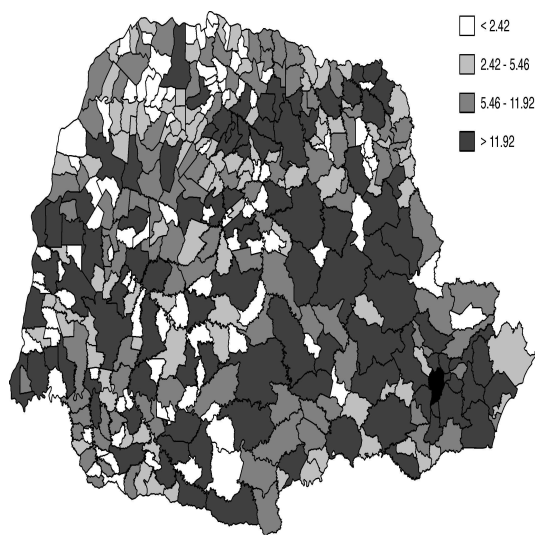


Figura 5.4: Média de óbitos por causas externas nos 26 anos estudados

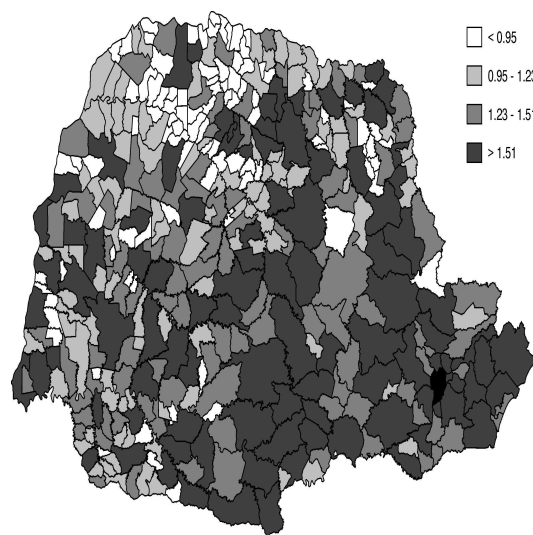


Figura 5.5: Média da covariável escola nos 26 anos estudados

Para analisar a evolução da correlação entre óbitos por causas externas e o logaritmo do número de escolas padronizadas, foram feitos mapas dessas duas variáveis nos anos de 1980, 1990, 1996 e 2004. Esses mapas são apresentados da Figura 5.6 até a Figura 5.13. Nas Figuras 5.6, 5.7 e 5.8 analisamos o logaritmo da covariável escola padronizada, na década de 80 e 90 os municípios com maior população aparecem com menor quantidade de escolas por habitante, para esta conclusão basta observar as áreas brancas no mapa. a partir de 2000, essa tendência continua, observar Figura 5.9.

Nas Figuras 5.10 e 5.11 apresentamos os mapas da variável óbitos por causas externas em 1980 e 1990 respectivamente. As regiões Metropolitanas de Curitiba e Foz do Iguaçu já apresentavam um elevado número de mortes, e devido ao aumento do tráfico de drogas na fronteira, o número de óbitos em Foz do Iguaçu e sua Região Metropolitana começou

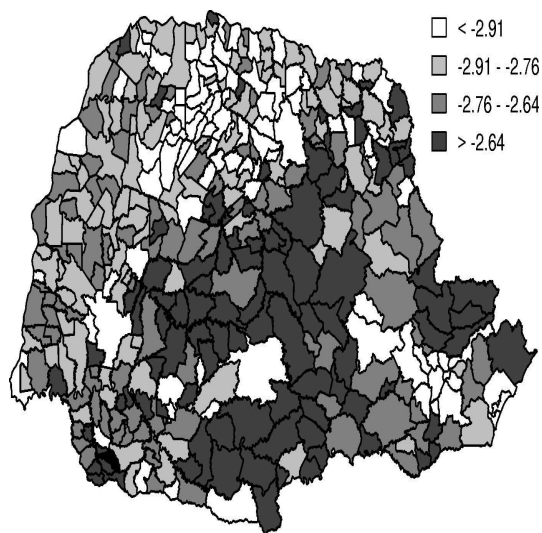


Figura 5.6: *Número de escolas / população*  
- 1980

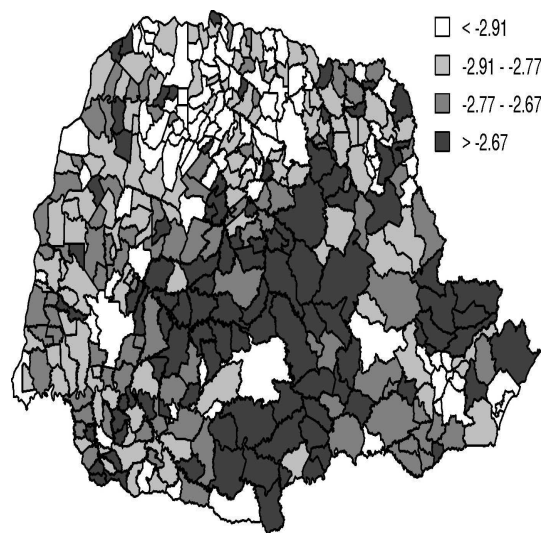


Figura 5.7: *Número de escolas / população*  
- 1990

aumentar. Os números observados deveriam ser ainda maiores, pois existe um grande número de subnotificações nessas áreas. Em uma análise mais recente, nos de 1996 e 2004, temos outras cidades como Londrina e Maringá começando a preocupar a Segurança Pública do Estado do Paraná. Os mapas do número de óbitos por causas externas em 1996 e 2004 estão nas Figuras 5.12 e 5.13.

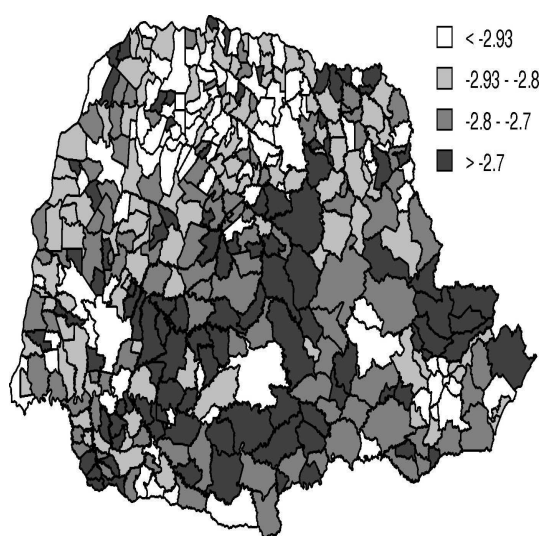


Figura 5.8: *Número de escolas / população*  
- 1996

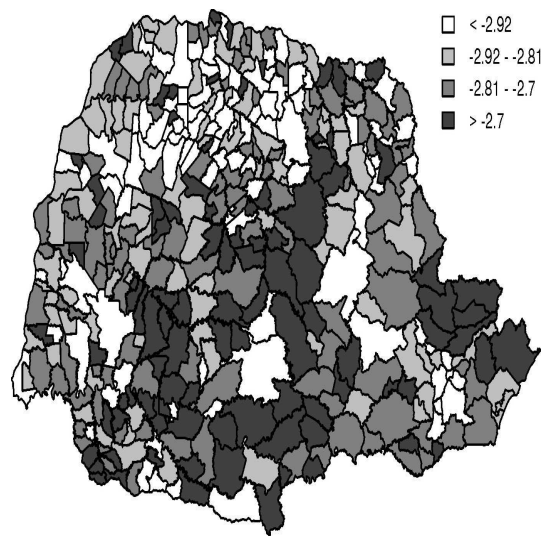


Figura 5.9: *Número de escolas / população*  
- 2004

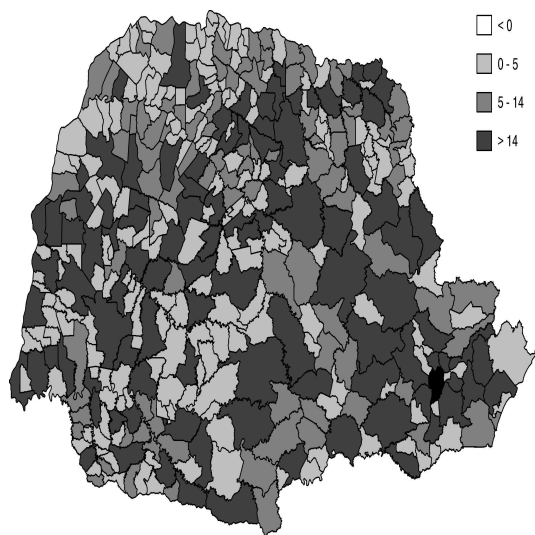


Figura 5.10: *Óbitos por causas externas no*  
*ano de 1980*

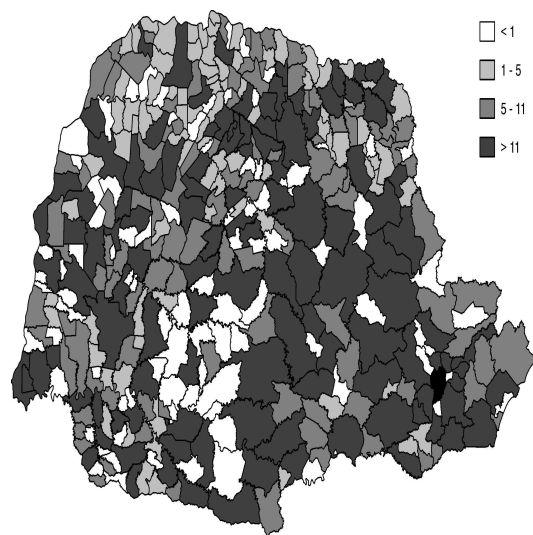


Figura 5.11: *Óbitos por causas externas no*  
*ano de 1990*

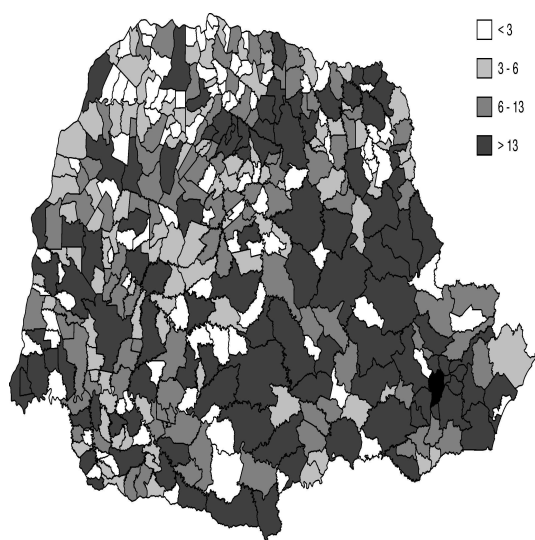


Figura 5.12: Óbitos por causas externas no ano de 1996

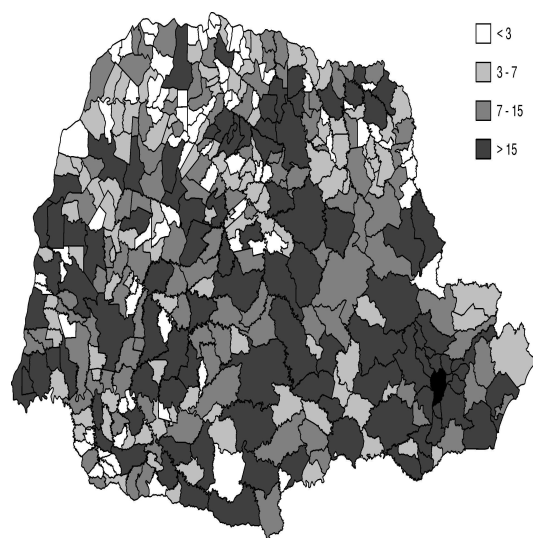


Figura 5.13: Óbitos por causas externas no ano de 2004

## 5.2 Cálculo do valor esperado de óbitos por município

Supondo que o número médio de óbitos por causas externas em determinada região é proporcional ao número de habitantes da mesma, e tendo observado o número total de óbitos no Estado do Paraná para um determinado período de tempo  $t$ , podemos calcular o número esperado de óbitos  $e_{it}$  em cada município  $i$  utilizando a seguinte expressão:  $e_{it} = p_{it}p_t$ , onde  $p_t = \frac{\sum_i y_{it}}{\sum_i p_{it}}$  e  $p_{it}$  é a população do município  $i$  no tempo  $t$ . Sob essa suposição, os municípios com maior população tem valores esperados maiores. Nas Figuras 5.14, 5.15, 5.16 e 5.17 temos o mapa de  $e_{it}$  nos anos de 1980, 1990, 1996 e 2004. Note que os  $e_{it}$ 's são proporcionais ao tamanho populacional do município  $i$  no tempo  $t$ .

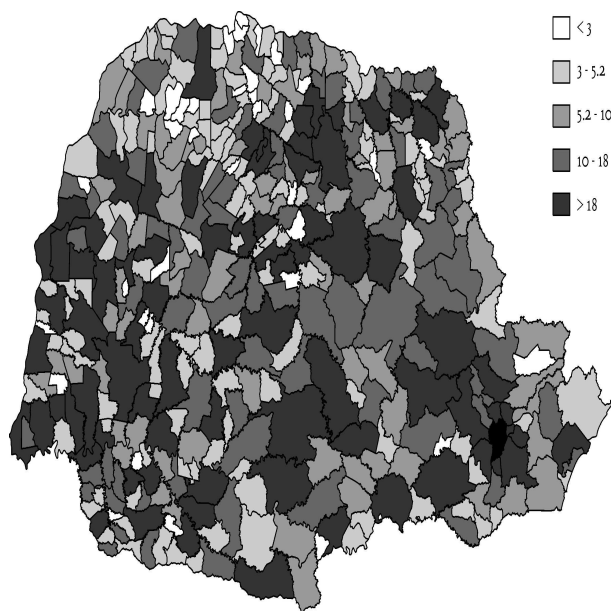


Figura 5.14: Valor esperado de óbitos por causas externas em 1980

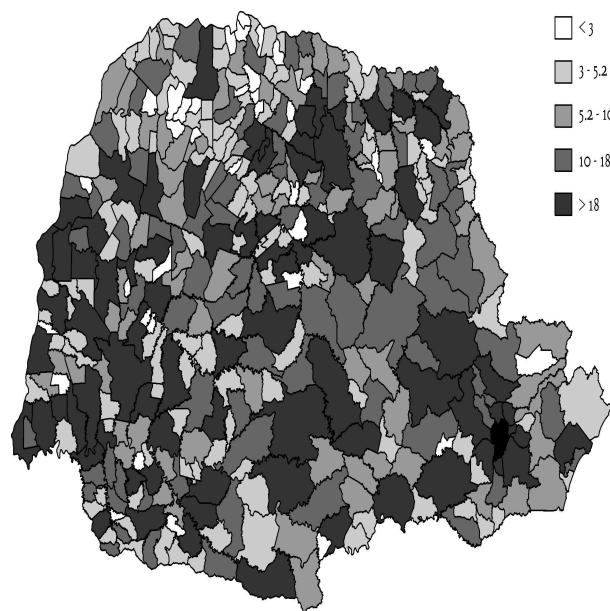


Figura 5.15: Valor esperado de óbitos por causas externas em 1990

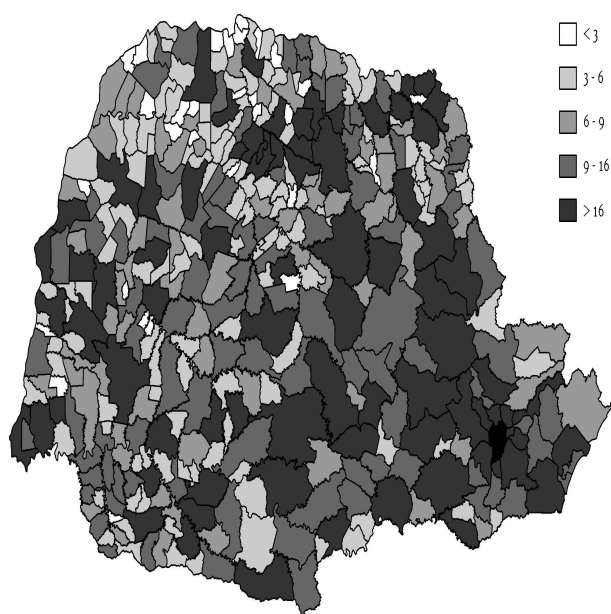


Figura 5.16: *Valor esperado de óbitos por causas externas em 1996*

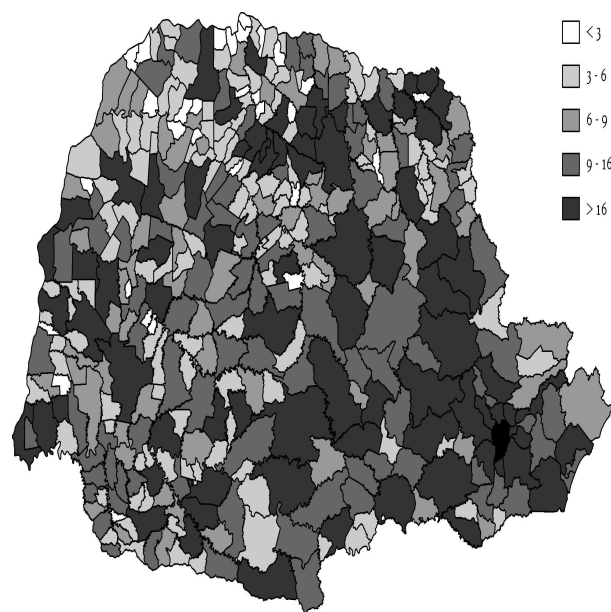


Figura 5.17: *Valor esperado de óbitos por causas externas em 2004*

As Figuras 5.18, 5.19, 5.20 e 5.21 mostram o gráfico dos valores observados de óbitos por causas externas ( $y_{it}$ ) versus os valores esperados de óbitos por causas externas ( $e_{it}$ ), nos 399 municípios do Estado do Paraná, para os anos de 1980, 1990, 1996 e 2004 respectivamente. Fica claro uma grande concordância dos valores destas duas variáveis, ou seja, a grande maioria dos municípios têm valores esperados de óbitos por causas externas próximos dos valores observados na realidade.

Fica nitido que existem indícios de uma tendência linear entre essas duas variáveis. Ajustamos o modelo de regressão  $y_{it} = \beta_0 + \beta_1 * e_{it} + \epsilon_{it}$ , onde  $\epsilon_{it} \sim N(0, \sigma^2)$ . Como o município de Curitiba pode estar influenciando muito na estimativa desse modelo, este foi estimado com e sem o ponto correspondente a Curitiba. Foi constatado que os parâmetros  $\beta_0$  e  $\beta_1$  são significativos em ambos os casos. Com essa análise verificamos que a população é altamente correlacionada com a variável óbitos por causas externas, como pode ser visto nas Figuras 5.22 e 5.23. Nosso interesse agora é investigar uma possível estrutura de dependência espacial nos resíduos da regressão, e uma possível correlação desses resíduos com a variável logaritmo do número de escolas por município.

Inicialmente, para análise dos resíduos ( $\epsilon_{it}$ ), apresentamos mapas incluindo Curitiba (Figura 5.24) e excluindo Curitiba (Figura 5.25) no ano de 2004. Em ambos os mapas podemos perceber uma estrutura espacial possivelmente significativa, ou seja, os municípios com resíduos grandes (com grande diferença entre o valor esperado e o observado) tendem a ser próximos de municípios com resíduos também grandes.

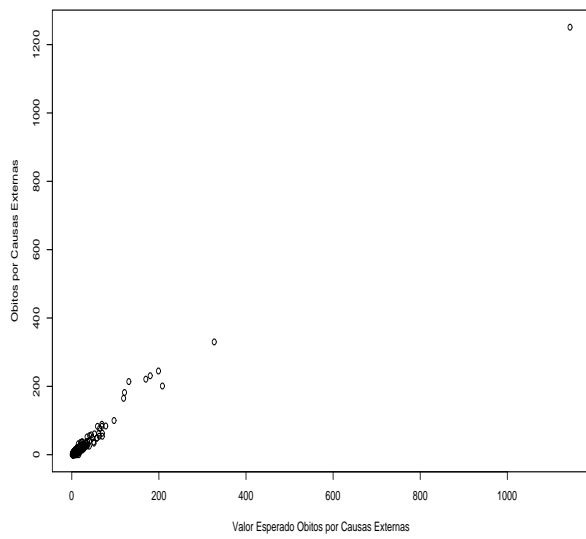


Figura 5.18: *Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1980*

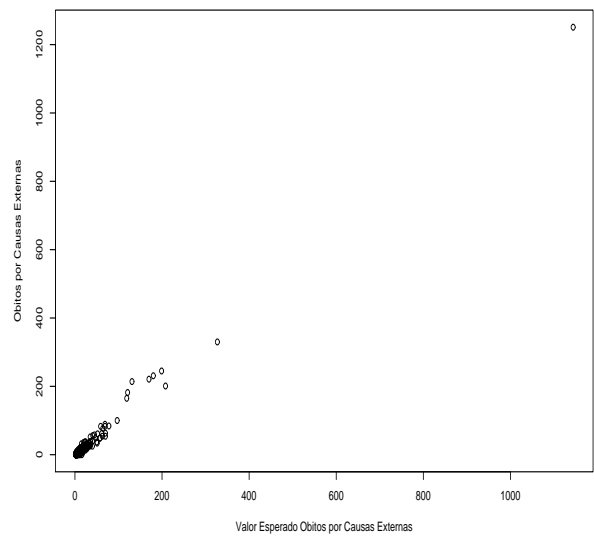


Figura 5.19: *Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1990*

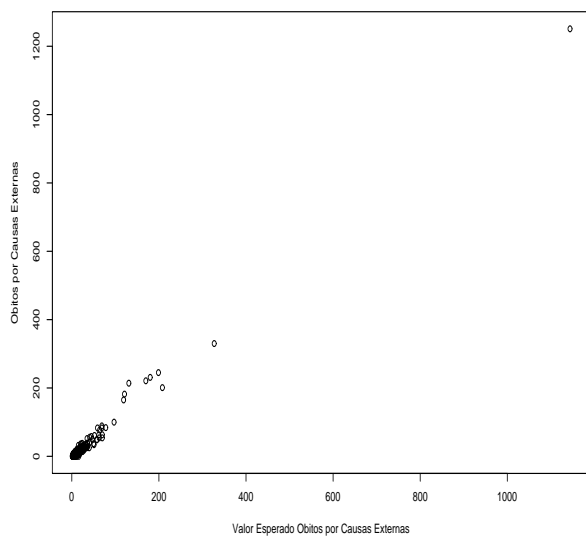


Figura 5.20: *Gráfico óbitos por causas externas versus seu respectivo valor esperado em 1996*

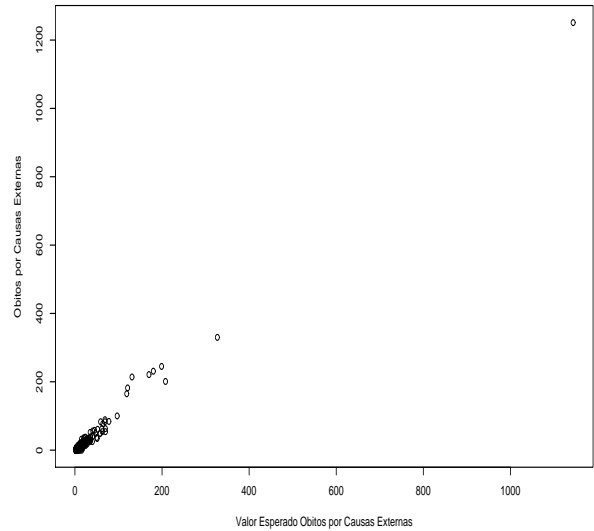


Figura 5.21: *Gráfico óbitos por causas externas versus seu respectivo valor esperado em 2004*



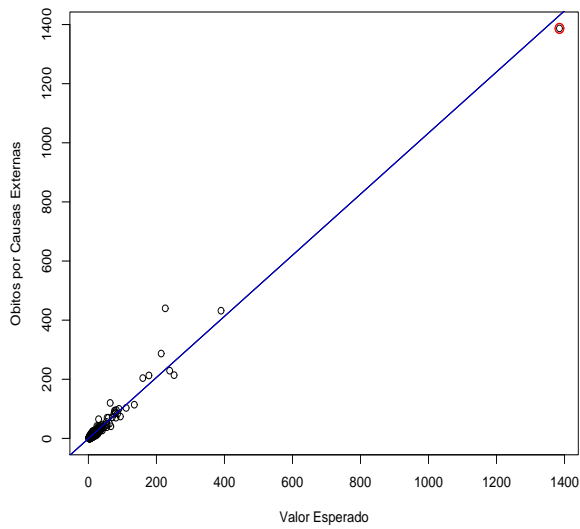


Figura 5.22: Gráfico óbitos por causas externas versus seu respectivo valor esperado com Curitiba em 2004

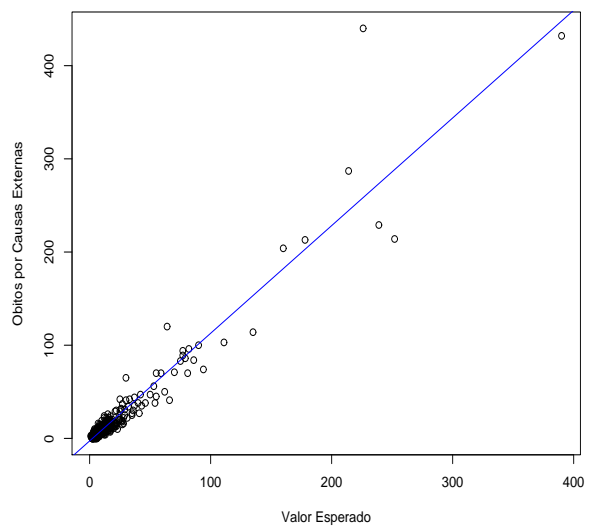


Figura 5.23: Gráfico óbitos por causas externas versus seu respectivo valor esperado sem Curitiba em 2004

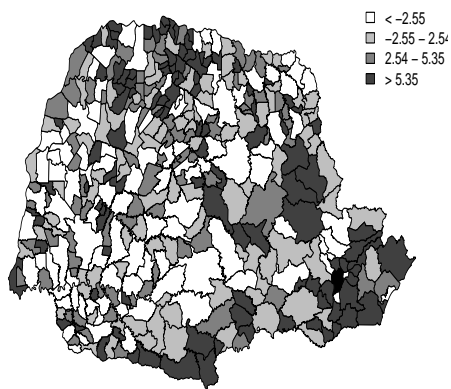


Figura 5.24: Resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado com Curitiba 2004

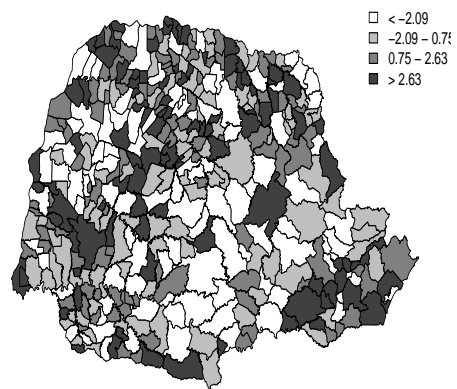


Figura 5.25: Resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado sem Curitiba 2004

Para tentar explicar a variação desses resíduos, estudamos a relação da covariável logaritmo do número de escolas padronizado pelo tamanho populacional e os resíduos do modelo ajustado anteriormente. Gráficos entre essas duas variáveis são apresentados na Figura 5.26 (para o ano de 2004 incluindo Curitiba), e Figura 5.27 (para o ano de 2004 excluindo Curitiba). Essas figuras indicam que não existe uma relação linear entre essas variáveis, porém os resíduos maiores em módulo tendem a aparecer em municípios onde temos menos escolas por habitantes.

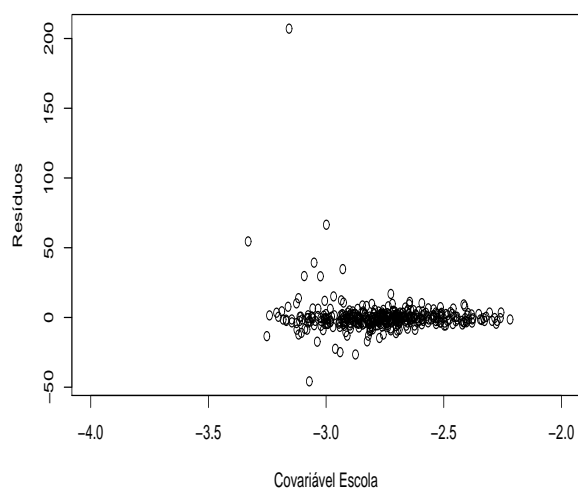


Figura 5.26: *Gráfico dos resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado com Curitiba versus logaritmo covariável escola padronizada em 2004.*

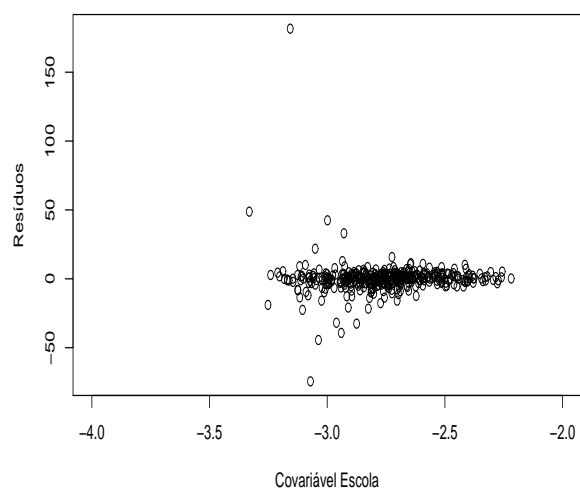


Figura 5.27: *Gráfico dos resíduos do modelo linear ajustado para óbitos por causas externas e seu respectivo valor esperado sem Curitiba versus logaritmo covariável escola padronizada em 2004.*

### 5.3 Cálculo do Standardized Mortality Ratio (SMR)

Alguns trabalhos de epidemiologia utilizam uma estimativa de máxima verossimilhança para o risco relativo. Assume-se que  $y_{it} \sim \text{Poisson}(\Lambda_{it})$ , onde  $\Lambda_{it} = \psi_{it}e_{it}$ ,

$i = 1, \dots, N$  e  $t = 1, \dots, T$ . Assume-se que o valor esperado  $e_{it}$ , definido anteriormente, é supostamente conhecido. A função de verossimilhança de  $\Lambda_{it}$  dado  $y_{it}$  é dada por:

$$L(\Lambda_{it}, y_{it}) = \frac{e^{-\Lambda_{it}} \Lambda_{it}^{y_{it}}}{y_{it}!}. \quad (5.1)$$

Aplicando o logaritmo nessa função, obtemos:

$$l(\Lambda_{it}, y_{it}) = \log(L(\Lambda_{it}, y_{it})) = -\Lambda_{it} + y_{it} \log(\Lambda_{it}) - \log(y_{it}!). \quad (5.2)$$

Verifica-se facilmente que o estimador de máxima verossimilhança de  $\Lambda_{it}$  é dado por  $\hat{\Lambda}_{it} = y_{it}$ ,  $i = 1, \dots, N$ ,  $t = 1, \dots, T$ .

Substituindo  $\Lambda_{it}$  por  $\psi_{it}e_{it}$ , temos que o estimador de máxima verossimilhança para  $\psi_{it}$  é dado por  $\hat{\psi}_{it} = \frac{y_{it}}{e_{it}}$ ,  $i = 1, \dots, N$ ,  $t = 1, \dots, T$ . Em vários trabalhos de mortalidade,  $\hat{\psi}_{it}$ , é denominado Standardized Mortality Ratio (SMR), ou seja,  $\hat{\psi}_{it}$  é uma taxa de mortalidade padronizada. A Figura 5.28 mostra o mapa do SMR para o ano de 2004.

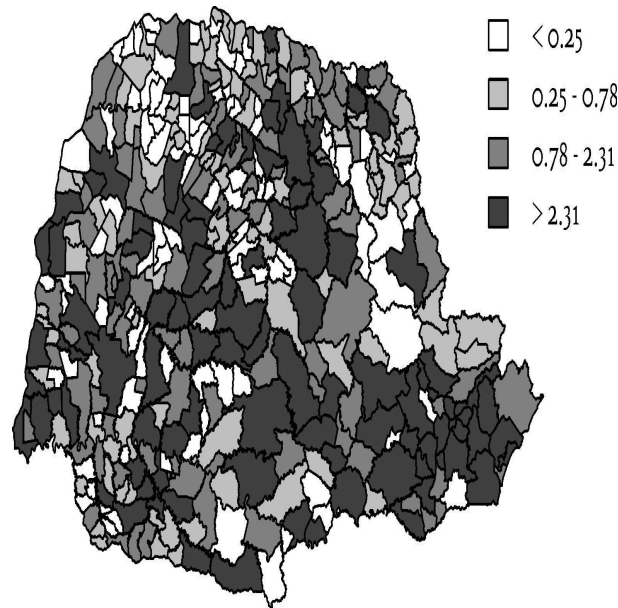


Figura 5.28: SMR óbitos por causas externas em 2004

## 5.4 Relação entre óbitos por causas externas e tráfico de drogas

Analisamos a relação entre as variáveis número de óbitos por causas externas e número de casos registrados de tráfico de drogas. Essa análise é feita apenas para o ano de 2004, visto que só temos informação sobre os casos de tráfico nesse ano. A Figura 5.29 mostra um gráfico de uma variável contra a outra. Através dessa figura fica visível que existe uma correlação positiva entre essas variáveis. A Figura 5.30 mostra esse mesmo gráfico, porém excluindo a cidade de Curitiba, para uma melhor visualização dos outros municípios. A correlação positiva entre as duas variáveis continua evidente.

A Figura 5.31 mostra o mapa do SMR do número de casos de tráfico de drogas. Comparando as Figuras 5.28 e 5.31, observamos uma correlação positiva entre o SMR para óbitos e tráfico, ou seja, municípios com alto SMR de tráfico possuem um alto SMR de óbitos por causas externas.

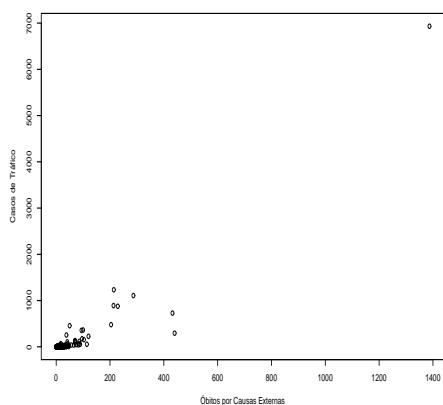


Figura 5.29: *Gráfico óbitos por causas externas versus número de casos de tráfico de drogas em 2004.*

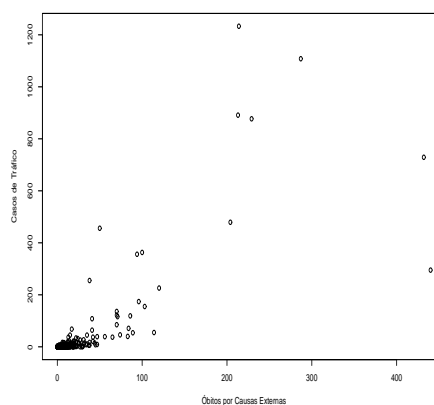


Figura 5.30: *Gráfico óbitos por causas externas versus número de casos de tráfico de drogas em 2004 sem Curitiba.*

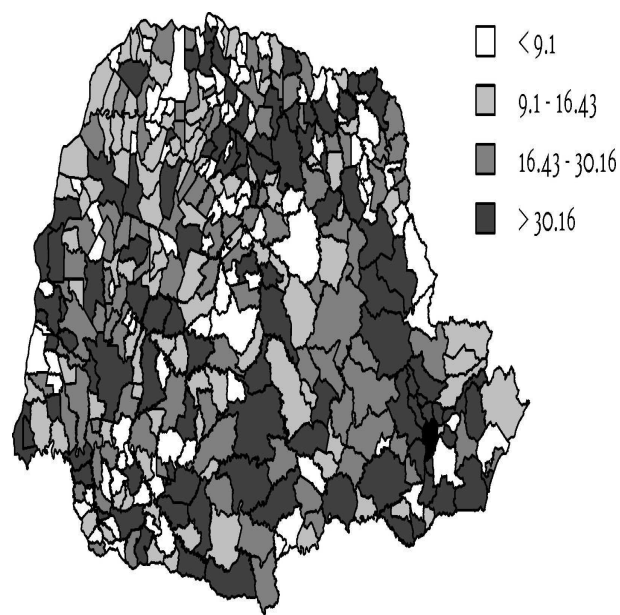


Figura 5.31: *SMR número de casos de tráfico de drogas em 2004.*

O capítulo a seguir serão abordados os modelos univariados espaço-temporais propostos para estimar o risco relativo, com resultados dos principais parâmetros para todos os modelos e seleção do melhor modelo através do critério DIC.

# Capítulo 6

## Modelagem Univariada

### 6.1 Especificação dos Modelos

Suponha uma região dividida em  $N$  regiões contíguas, onde contagens de um evento de interesse são feitas em instantes sucessivos de tempo. Seja  $y_{it}$  a variável de interesse (contagem) observada na região  $i$ ,  $i = 1, 2, \dots, N$  e no período de tempo  $t$ ,  $t = 1, 2, \dots, T$ .

Em geral dados de contagem são modelados através de uma distribuição de Poisson. É proposto que o número de óbitos por causas externas ( $y_{it}$ ) tenha distribuição de Poisson, com valor esperado dado pela multiplicação de dois componentes:  $e_{it}$ , que é uma intensidade supostamente conhecida que corresponde ao valor esperado de óbitos do município  $i$  no tempo  $t$  e  $\psi_{it}$ , que representa o risco relativo para  $i = 1, \dots, N$ ,  $t = 1, \dots, T$ :

$$y_{it} \sim Poi(\Lambda_{it})$$

$$\Lambda_{it} = e_{it}\psi_{it}$$

Logo,  $\log(e_{it}\psi_{it}) = \log(e_{it}) + \log(\psi_{it})$ , desta maneira é separado o efeito da população representado por  $\log(e_{it})$  do  $\log(\psi_{it})$ .

A distribuição proposta para o risco relativo  $\psi_{it}$  deve ser estruturada para acomodar os efeitos dos fatores dentro de cada área, além da variação espacial. Para ajudar a

explicar a variação de  $\psi_{it}$  é utilizado o logaritmo do número de escolas divididas pela sua respectiva população como variável explicativa (denotamos por  $X_{it}$ ).

A proposta da literatura é modelar o logaritmo do risco relativo como propõem Nobre, Schimdt e Lopes (2005) e Ferreira e Schimdt (2006) e é isso que será feito nesse trabalho. Outras referências para modelagem espaço-temporal são Kypraios (2003), Paez(2000) e Paez(2004). Nosso objetivo é propor diferentes estruturas para o logaritmo do risco relativo  $\log(\psi_{it})$ , atribuindo prioris adequadas para os parâmetros e comparar os resultados. Em todas essas propostas,  $\log(\psi_{it})$  é escrito como soma de diferentes componentes. Esses modelos, com suas respectivas prioris, serão dados a seguir:

- Modelo 1 (M1):  $\log(\psi_{it}) = \alpha + \beta X_{it} + \theta_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\alpha$  é o intercepto da série,  $\beta$  é o coeficiente de regressão e  $\theta_{it} \sim N(0, \sigma_t^2)$ ,  $i = 1, \dots, N$ , é um componente estocástico não estruturado, onde  $N(\mu, \sigma^2)$  denota a distribuição Normal com média  $\mu$  e variância  $\sigma^2$ . Assumimos para o intercepto e o coeficiente de regressão as seguintes distribuições a priori respectivamente  $\alpha \sim N(0, \sigma_\alpha^2)$  e  $\beta \sim N(0, \sigma_\beta^2)$
- Modelo 2 (M2):  $\log(\psi_{it}) = \alpha_t + \beta X_{it} + \theta_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\alpha_t$  é o intercepto da série no tempo  $t$ ,  $\beta$  é o coeficiente de regressão. e  $\theta_{it} \sim N(0, \sigma_t^2)$ ,  $i = 1, \dots, N$ , são componentes estocásticos não estruturados. Assumimos a priori que o intercepto  $\alpha_t$  segue um passeio aleatório:  $\alpha_t = \alpha_{t-1} + w_\alpha$  onde  $w_\alpha \sim N(0, \sigma_\alpha^2)$  e  $\beta \sim N(0, \sigma_\beta^2)$
- Modelo 3 (M3):  $\log(\psi_{it}) = \alpha_t + \beta_t X_{it} + \theta_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\alpha_t$  é o intercepto da série no tempo  $t$ ,  $\beta_t$  é o coeficiente de regressão no tempo  $t$  e  $\theta_{it} \sim N(0, \sigma_t^2)$ ,  $i = 1, \dots, N$ , são componentes estocásticos não estruturados. Assumimos a priori que os coeficientes de regressão variam no tempo suavemente de acordo com um passeio aleatório:  $\beta_t = \beta_{t-1} + w_t$  onde  $w_t \sim N(0, \sigma_w^2)$ . Assumimos também a priori que o intercepto  $\alpha_t$  varia suavemente no tempo, como foi definido no modelo 2.
- Modelo 4 (M4):  $\log(\psi_{it}) = \alpha + \beta X_{it} + \phi_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\phi'_t = (\phi_{1t}, \dots, \phi_{Nt})$ ,  $\phi_t \sim CAR(\sigma_t^2)$  é um componente estocástico estruturado no



espaço,  $\beta$  é definido como no Modelo 1 e para  $\alpha$  é necessário impor a priori uma priori Uniforme no intervalo  $(-\infty, \infty)$ , como foi comentado no capítulo 3.

- Modelo 5 (M5):  $\log(\psi_{it}) = \alpha_t + \beta X_{it} + \phi_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\phi'_t = (\phi_{1t}, \dots, \phi_{Nt})$ ,  $\phi_{it} \sim CAR(\sigma_t^2)$  é um componente estocástico estruturado no espaço,  $\beta$  é definido como no Modelo 1,  $\alpha_t$  já foi comentado no Modelo 4.
- Modelo 6 (M6):  $\log(\psi_{it}) = \alpha_t + \beta_t X_{it} + \phi_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\phi'_t = (\phi_{1t}, \dots, \phi_{Nt})$ ,  $\phi_{it} \sim CAR(\sigma_t^2)$  é um componente estocástico estruturado no espaço,  $\beta$  é definido como no Modelo 3,  $\alpha_t$  já foi comentado no Modelo 4.
- Modelo 7 (M7):  $\log(\psi_{it}) = \alpha_t + \phi_{it}$ , para  $t = 1, \dots, T$  e  $i = 1, \dots, N$ , onde  $\phi'_t = (\phi_{1t}, \dots, \phi_{Nt})$ ,  $\phi_{it} \sim CAR(\sigma_t^2)$  é um componente estocástico estruturado no espaço, como no Modelo 3, para  $\alpha_t$  é necessário impor uma priori Uniforme no intervalo  $(-\infty, \infty)$ , como foi comentado no capítulo 3.

Para modelar  $\sigma_t^2$  utilizamos duas propostas: na primeira, os  $\sigma_t^2$ 's são independentes e identicamente distribuídos com distribuição a priori Gama Invertida com média  $a/b$  e variância  $a/b^2$  (que denotaremos por  $\sigma_t^2 \sim IG(a, b)$ ), na segunda, propomos que  $\log(\sigma_t^2)$  siga um passeio aleatório, de acordo com a equação  $\log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \epsilon_t$ , onde  $\epsilon_t \sim N(0, \tau_\epsilon^2)$ .

Resumindo, nos modelos 1, 2 e 3 está sendo assumido que o componente estocástico é independente para todos os municípios durante todos os anos. No modelo 1 o intercepto e o coeficiente de regressão não variam no tempo, no modelo 2 apenas o intercepto varia no tempo e no modelo 3 ambos parâmetros variam.

Nos modelos 4, 5 e 6 incorporamos ao componente estocástico uma estrutura espacial. No modelo 4 o intercepto e o coeficiente de regressão não variam no tempo, no modelo 5 apenas o intercepto varia no tempo e no modelo 6 ambos parâmetros variam. O modelo 7 é similar ao modelo 5, porém excluindo a variável explicativa  $X_{it}$ .

## 6.2 Inferência

Com a especificação dos modelos e a escolha das distribuições a priori, nosso objetivo agora é a obtenção das distribuições a posteriori dos parâmetros desconhecidos de cada modelo. Essas distribuições podem ser otidas pela abordagem Bayesiana através do teorema de Bayes, serão apresentadas a seguir.

- Inferência sob o Modelo 1

Seja  $\varphi$  a coleção de todos os parâmetros desconhecidos no Modelo 1, ou seja,  $\varphi = (\alpha, \beta, \theta, \sigma^2, \sigma_b^2, \sigma_\alpha^2)$ , onde,  $\theta = (\theta_1, \dots, \theta_T)^T$ ,  $\theta_t = (\theta_{1t}, \dots, \theta_{Nt})^T$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Para obter a distribuição a posteriori  $p(\varphi|y)$  utilizamos o Teorema de Bayes, e temos o seguinte resultado:

$$\begin{aligned} p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\ p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha, \beta, \theta_t, \sigma_t^2, \sigma_b^2, \sigma_\alpha^2) \\ p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} p(\alpha|\sigma_\alpha^2)p(\beta|\sigma_b^2) \prod_{t=1}^T p(\theta_t|\sigma_t^2)p(\sigma_t^2) \end{aligned}$$

- Inferência sob o Modelo 2

Vamos agora definir  $\varphi = (\alpha, \beta, \theta, \sigma^2, \sigma_b^2, \sigma_\alpha^2)$ , onde  $\alpha = (\alpha_1, \dots, \alpha_T)^T$ ,  $\theta = (\theta_1, \dots, \theta_T)^T$ ,  $\theta_t = (\theta_{1t}, \dots, \theta_{Nt})$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de Bayes, temos que:

$$\begin{aligned} p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\ p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t, \beta, \theta_t, \sigma_t^2, \sigma_b^2) \\ p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} p(\beta|\sigma_b^2) \prod_{t=1}^T p(\alpha_t|\alpha_{t-1}, \sigma_\alpha^2)p(\theta_t|\sigma_t^2)p(\sigma_t^2) \end{aligned}$$

- Inferência sob o Modelo 3

Vamos agora definir  $\varphi = (\alpha, \beta, \theta, \sigma^2, \sigma_b^2, \sigma_\alpha^2)$ , onde  $\alpha = (\alpha_1, \dots, \alpha_T)^T$ ,  $\beta = (\beta_1, \dots, \beta_T)^T$ ,  $\theta = (\theta_1, \dots, \theta_T)^T$ ,  $\theta_t = (\theta_{1t}, \dots, \theta_{Nt})$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de Bayes

temos que:

$$\begin{aligned}
p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t, \beta_t, \theta_t, \sigma_t^2, \sigma_b^2, \sigma_\alpha^2) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t|\alpha_{t-1}, \sigma_\alpha^2) p(\beta_t|\beta_{t-1}, \sigma_b^2) p(\theta_t|\sigma_t^2) p(\sigma_t^2)
\end{aligned}$$

- Inferência sob o Modelo 4

Vamos agora definir  $\varphi = (\alpha, \beta, \phi, \sigma^2, \sigma_b^2)$ ,  $\phi = (\phi_1, \dots, \phi_T)^T$ ,  $\phi_t = (\phi_{1t}, \dots, \phi_{Nt})^T$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de Bayes temos que:

$$\begin{aligned}
p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha, \beta, \phi_t, \sigma_t^2, \sigma_b^2) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} p(\alpha) p(\beta|\sigma_b^2) \prod_{t=1}^T p(\phi_t, |\sigma_t^2) p(\sigma_t^2)
\end{aligned}$$

- Inferência sob o Modelo 5

Vamos agora definir  $\varphi = (\alpha, \beta, \phi, \sigma^2, \sigma_b^2)$ , onde  $\alpha = (\alpha_1, \dots, \alpha_T)^T$ ,  $\phi = (\phi_1, \dots, \phi_T)^T$ ,  $\phi_t = (\phi_{1t}, \dots, \phi_{Nt})^T$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de Bayes temos que:

$$\begin{aligned}
p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t, \beta, \phi_t, \sigma_t^2, \sigma_b^2) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} p(\beta|\sigma_b^2) \prod_{t=1}^T p(\alpha_t) p(\phi_t|\sigma_t^2) p(\sigma_t^2)
\end{aligned}$$

- Inferência sob o Modelo 6

Vamos agora definir  $\varphi = (\alpha, \beta, \phi, \sigma^2, \sigma_b^2)$ , onde  $\alpha = (\alpha_1, \dots, \alpha_T)^T$ ,  $\beta = (\beta_1, \dots, \beta_T)^T$ ,  $\phi = (\phi_1, \dots, \phi_T)^T$ ,  $\phi_t = (\phi_{1t}, \dots, \phi_{Nt})^T$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de

Bayes temos que:

$$\begin{aligned}
p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t, \beta_t, \phi_t, \sigma_t^2, \sigma_b^2) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t)p(\beta_t|\beta_{t-1}, \sigma_b^2)p(\phi_t|\sigma_t^2)p(\sigma_t^2)
\end{aligned}$$

- Inferência sob o Modelo 7

Vamos agora definir  $\varphi = (\alpha, \phi, \sigma^2)$ , onde  $\alpha = (\alpha_1, \dots, \alpha_T)^T$ ,  $\phi_t = (\phi_{1t}, \dots, \phi_{Nt})^T$  e  $\sigma^2 = (\sigma_1^2, \dots, \sigma_T^2)^T$ . Pelo Teorema de Bayes temos que:

$$\begin{aligned}
p(\varphi|y) &\propto p(y|\varphi)p(\varphi) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t, \phi_t, \sigma_t^2) \\
p(\varphi|y) &\propto \prod_{i=1}^N \prod_{t=1}^T \exp \{y_{it} \log(e_{it}\psi_{it}) - e_{it}\psi_{it} - \log(y_{it}!)\} \prod_{t=1}^T p(\alpha_t)p(\phi_t|\sigma_t^2)p(\sigma_t^2)
\end{aligned}$$

As distribuições a posteriori obtidas sob os modelos apresentados acima não possuem forma conhecida. Dessa forma, é necessário a utilização de métodos computacionalmente intensivos para fazer inferência a respeito dos parâmetros desconhecidos desses modelos. Neste trabalho vamos utilizar o método MCMC, que foi apresentado no capítulo 4. O software utilizado para estimação dos parâmetros por MCMC foi o WinBugs (Bayesian Using Gibbs), implementado por Thomas et al. (1992). Este pacote estatístico já tem implementado internamente as rotinas para estimar os parâmetros via MCMC, bastando apenas fazer a especificação do modelo e das distribuições a priori dos parâmetros. O amostrador de Gibbs implementado no WinBugs utiliza duas rotinas diferentes: ARS (Adaptive Rejection Sampling) e ARMS (Adaptive Rejection Metropolis Sampling). A primeira pode ser usada para amostrar de forma eficiente qualquer distribuição condicional cuja função densidade é log-côncava. A segunda generaliza a rotina ARS para o caso de funções que não sejam log-côncavas. Caso não seja possível construir as distribuições condicionais completas, o software utiliza o algoritmo Metropolis-Hastings onde

a distribuição proposta  $q(., \theta^t)$  é uma distribuição gaussiana centrada no valor atual do parâmetro  $\theta^t$ . No apêndice está especificado o código para realizar inferência no programa WinBugs

### 6.3 Critério *DIC* (Deviance Information Criterion) para Comparação de Modelos

Para a comparação dos modelos, utilizaremos a estatística *DIC* (Spiegelhalter et. al., 2002), que favorece modelos com melhor bondade de ajuste e penaliza pelo excesso de parâmetros.

Seja  $\hat{\varphi}$  o vetor de médias a posteriori dos parâmetros  $\varphi$  sob determinado modelo. Vamos definir a função desvio por  $D = -2\log(p(y|\varphi))$ . Para calcular o *DIC*, devemos primeiramente calcular a esperança a posteriori da estatística  $D$ . O passo seguinte é calcular o desvio fazendo  $\varphi = \hat{\varphi}$ . Definimos assim as seguintes quantidades:

$$\begin{aligned} D^* &= E_{\varphi|y}[D], \\ D^{**} &= -2\log(p(y|\hat{\varphi})). \end{aligned}$$

Uma forma de mensurar a complexidade do modelo, ou seja, verificar o número de parâmetros efetivos envolvidos no modelo, é através da seguinte fórmula:

$$pD = D^* - D^{**}.$$

A partir das equações acima podemos definir o *DIC* como sendo:

$$DIC = D^* + pD = D^{**} + 2pD.$$

O melhor modelo de acordo com a estatística *DIC* é aquele onde essa estatística apresenta menor valor. O ponto negativo do *DIC* é que caso a média a posteriori não seja uma boa estimativa os resultados se tornam não confiáveis.

## 6.4 Resultados Obtidos

Nessa subseção apresentaremos os resultados obtidos com a aplicação dos sete modelos univariados propostos ao número de óbitos por causas externas no estado do Paraná. Para cada modelo, obtivemos resultados considerando duas prioris diferentes para  $\sigma_t$ , que foram descritas anteriormente: a primeira assumindo que os  $\sigma_t$ 's são independentes (proposta de P1); e a segunda assumindo que  $\sigma_t$  segue um passeio aleatório (proposta de P2). Como as distribuições a posteriori dos modelos de interesse não tem forma analítica fechada, a estimação dos parâmetros foi feita por MCMC e para isso foi utilizado o software WinBugs. O período de aquecimento das cadeias, “burn-in”, foi de 50000 iterações. Após o “burn-in” foram geradas mais 50000 iterações, que foram utilizadas para fazer inferência. A convergência das cadeias foi observada através de uma inspeção visual. Na Tabela 6.1, apresentamos o valor do DIC obtido em cada um dos 14 modelos. O melhor modelo segundo esse critério é aquele que apresenta menor valor de DIC. Pela Tabela 6.1, observamos que o modelo com menor valor de DIC foi o Modelo 7 (intercepto variando no tempo e sem a covariável), com a proposta 2 (passeio aleatório para a variância do componente  $\phi_{it}, i = 1, \dots, N$ ). Como o fator populacional representado pelo valor esperado, explica em grande parte a distribuição de óbitos por causas externas no estado do Paraná, não foi necessário incorporar a covariável no modelo.

Valores DIC		
Modelos	P1	P2
M1	52887.600	54701.100
M2	52404.600	52339.250
M3	52376.400	53169.400
M4	54734.400	54703.500
M5	52962.200	53259.150
M6	52911.600	53043.300
M7	52399.210	52329.240

Tabela 6.1: Tabela DIC dos 14 modelos univariados

Para os 14 modelos analisados, em muitas situações o comportamento dos parâmetros apresentaram resultados semelhantes. A seguir vamos mostrar uma sequência de resultados dos principais parâmetros de alguns desses modelos.

### Modelos com $\alpha$ e $\beta$ fixos

Na Tabela 6.2 apresentamos a média a posteriori dos modelos onde  $\alpha$  e  $\beta$  foram assumidos fixos, ou seja,  $M_1$  e  $M_4$ . Percebemos que os modelos com efeitos estocásticos estruturados e não estruturados quase não se diferem na estimativa desses parâmetros. Também podemos perceber que a escolha da propostas P1 ou P2 para a variância do componente estocástico não altera sensivelmente os resultados. Em ambos os modelos à covariável foi significativa e com sinal negativo, indicando que o aumento da quantidade de escolas por habitante resulta em um menor risco relativo.

Média a Posteriori				
Parâmetros	M1 com P1	M1 com P2	M4 com P1	M4 com P2
$\alpha$	-0.268	-0.2309	-0.25	-0.2556
$\beta$	-0.210	-0.2200	-0.21	-0.2200

Tabela 6.2: Média a Posteriori

Na Tabela 6.3 analisamos o desvio padrão a posteriori dos parâmetros  $\alpha$  e  $\beta$ . Quando construímos o intervalo de credibilidade percebemos que todos os parâmetros são significativos pois o zero não está contido nos intervalos de credibilidade.

Desvio Padrão a Posteriori				
Parâmetros	M1 com P1	M1 com P2	M4 com P1	M4 com P2
$\alpha$	0.016	0.011	0.02	0.02
$\beta$	0.011	0.06	0.013	0.06

Tabela 6.3: Desvio Padrão a Posteriori

### Análise do parâmetro $\alpha_t$ variando no tempo

Em todos os modelos estudados com  $\alpha_t$  variando no tempo, os resultados de inferência para  $\alpha_t$  foram semelhantes. As Figuras 6.1 e 6.2 mostram os intervalos de credibilidade para  $\alpha_t$  para cada período de tempo  $t$  sob o modelo  $M5$  com as propostas P1 e P2 para

a variância  $\sigma_t^2$ . Constatamos por essas figuras uma variação suave ao longo do tempo, justificando utilizar modelos dinâmicos para modelar  $\alpha_t$ . Observamos também que nos últimos anos os valores de  $\alpha_t$  começam a ter uma menor variabilidade, implicando que os riscos relativos sofrem menor variação.

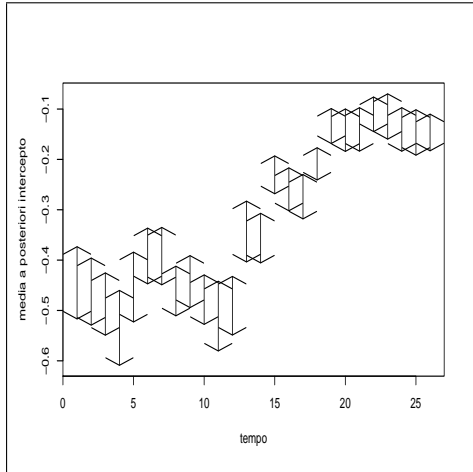


Figura 6.1: *Gráfico de máxima densidade a posteriori do Intercepto do M5 com P1 nos 26 anos estudados*

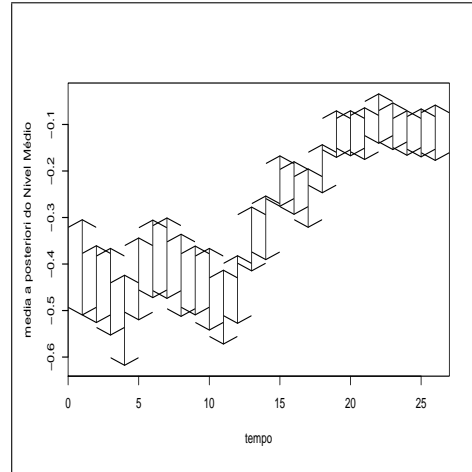


Figura 6.2: *Gráfico de máxima densidade a posteriori do Intercepto do M5 com P2 nos 26 anos estudados*

### **Análise do parâmetro $\sigma_t^2$**

Ao invés de analisarmos a variância  $\sigma_t^2$ , escolhemos trabalhar com o seu inverso,  $\sigma_t^{-2}$ , que chamamos de precisão. Uma característica encontrada de  $\sigma_t^{-2}$  em todos os modelos foi a existência de dois regimes distintos: uma precisão menor até 1990 e após isso um aumento. Para ilustrar este comportamento, mostramos nas Figuras 6.3 e 6.4 os intervalos de credibilidade a posteriori da precisão ao longo do tempo para os modelos M6 com P1 e P2. Este aumento da precisão implica na diminuição da variabilidade do risco relativo. Essa redução da variabilidade pode ser devido à uma melhor qualidade de registros em hospitais, levando a um banco de dados mais consistente do DataSus.

A seguir serão apresentados alguns resultados de inferência para os principais parâmetros do Modelo 7 com proposta 2 para variância, escolhido como melhor modelo pelo critério



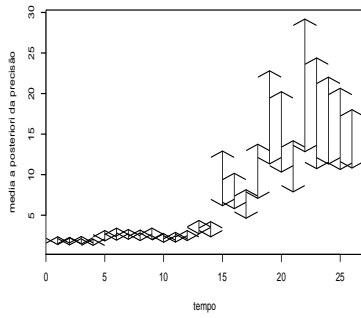


Figura 6.3: *Gráfico dos intervalos de máxima densidade a posteriori da precisão do M6 com P1 nos 26 anos estudados*

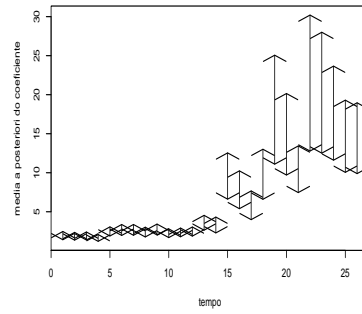


Figura 6.4: *Gráfico dos intervalos de máxima densidade a posteriori da precisão do M6 com P2 nos 26 anos estudados*

DIC. Na Figura 6.5 construímos os intervalos de credibilidade da precisão , ao longo do tempo. Novamente observamos dois regimes distintos na evolução da precisão - a partir de 1990 percebemos uma acentuada queda na variabilidade (ou aumento na precisão), que pôde ser capitada pela proposta 2 para  $\sigma_t$ . Na Figura 6.6 acompanhamos a trajetória da média a posteriori do intercepto ao longo do tempo através de intervalos de credibilidade da posteriori. Percebemos um crescimento suave neste período.

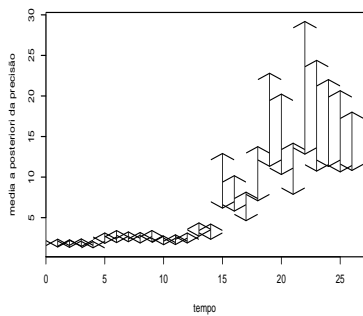


Figura 6.5: *Gráfico dos intervalos de máxima densidade a posteriori da precisão do modelo selecionado nos 26 anos estudados*

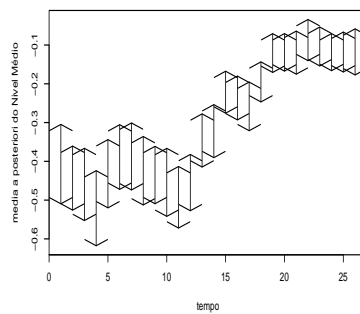


Figura 6.6: *Gráfico dos intervalos de máxima densidade a posteriori do intercepto do modelo selecionado nos 26 anos estudados*

Nas Figura 6.7 à 6.14, comparamos duas maneiras distintas de calcular o risco relativo: através do  $SMR_{it} = \frac{y_{it}}{e_{it}}$  e através de estimativas do risco relativo sob o enfoque Bayesiano (considerando o modelo escolhido pelo critério DIC). As figuras mostram gráficos do risco relativo sob os dois métodos versus o logaritmo da população nos anos de 1980, 1990, 1996 e 2004. Quando calculamos o  $SMR$  é feita uma razão entre o valor observado e o esperado. Vale lembrar que os municípios tem o valor esperado de óbito proporcional ao tamanho de sua população. Caso seja um município pouco populoso, e exista uma situação atípica, por exemplo um grave acidente automobilístico com uma quantidade de óbitos expressiva, o  $SMR$  fica inflacionado.

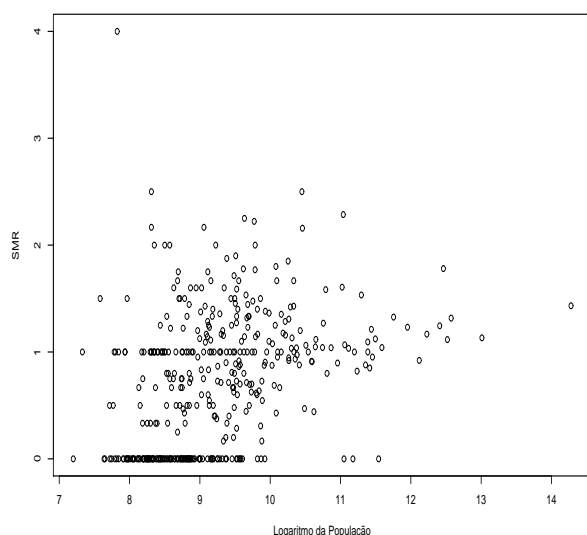


Figura 6.7: *Gráfico SMR versus logaritmo da população em 1980*

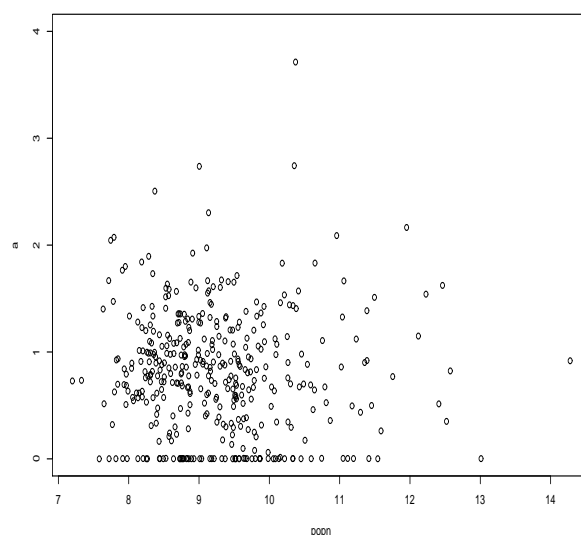


Figura 6.8: *Média a posteriori versus logaritmo da população em 1980*

Um dos objetivos da nossa proposta através da metodologia Bayesiana é corrigir superestimativas e subestimativas de riscos relativos para municípios com pequenas populações. Através das figuras podemos perceber que o  $SMR$  dos municípios com menor porte populacional tem uma grande variabilidade, mostrando que essas estimativas são provavelmente influenciadas por casos atípicos. As estimativas obtidas através do modelo proposto sob enfoque Bayesiano mostram a redução dessa variabilidade, minimizando o problema das sub e superestimativas encontradas em municípios com pequenas

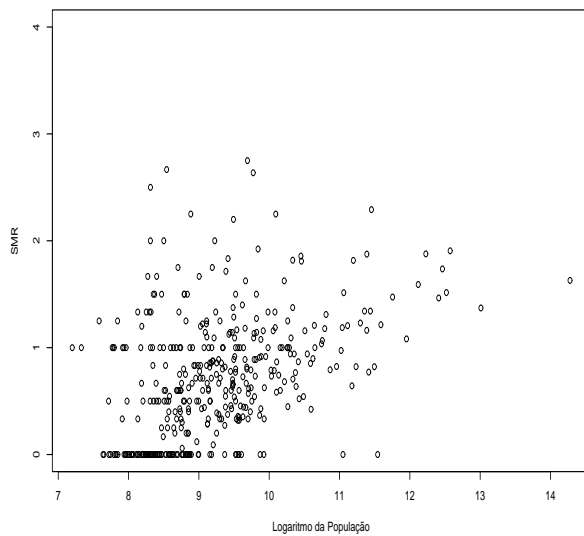


Figura 6.9: *SMR versus logaritmo da população em 1990*

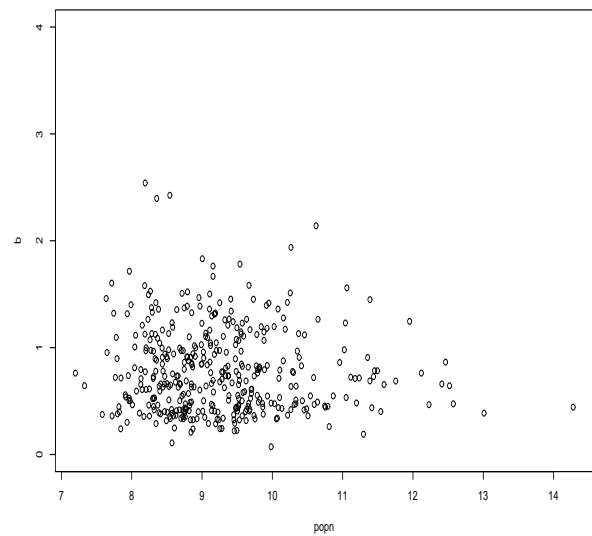


Figura 6.10: *Média a posteriori versus logaritmo da população em 1990*

populações.

Particularmente, no ano de 2004 o *SMR* do município de Figueira foi de aproximadamente 3, um dos maiores encontrados. Na prática esta cidade, que tem população inferior a 10000 habitantes, não requer cuidados especiais. Através do modelo Bayesiano proposto, o risco relativo estimado para o município de Figueira foi de aproximadamente 1, retratando melhor a realidade. Um ponto importante a ressaltar é que em municípios com população inferiores a 10000 habitantes, os homicídios são de natureza passional e acontecem em bares devido ao consumo exagerado de álcool, o que é diferente da dinâmica dos grandes municípios. Nos grandes municípios, os óbitos acontecem devido a situações sociais e econômicas e também no trânsito (acidentes podem ser enquadrados na lei como homicídio doloso). Podemos destacar a cidade de Curitiba, cuja frota de carros triplicou nos últimos 10 anos.

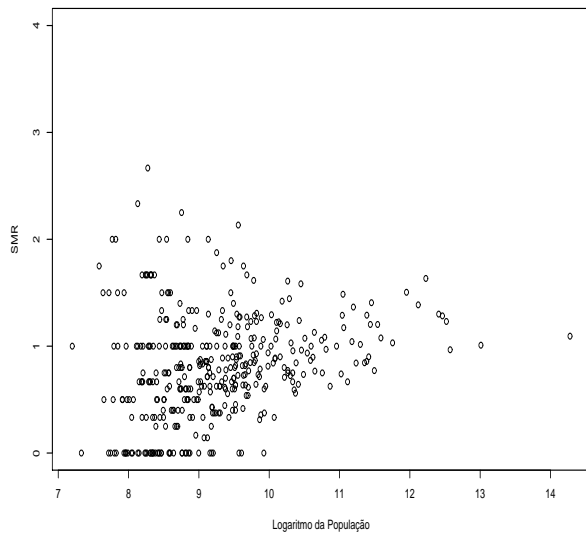


Figura 6.11: *SMR versus logaritmo da população 1996*

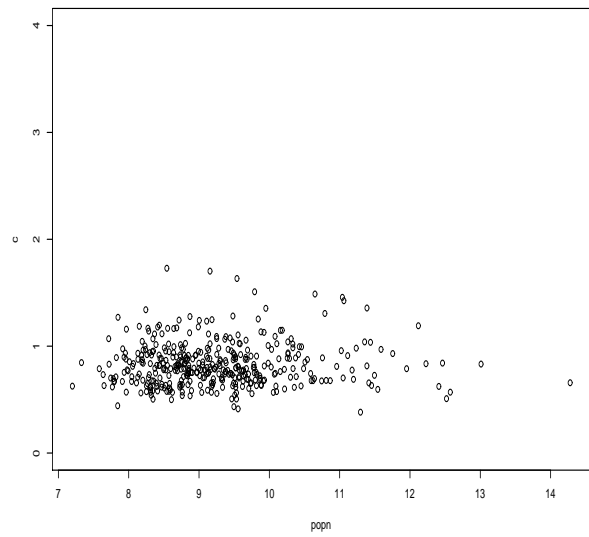


Figura 6.12: *Média a posteriori versus logaritmo da população em 1996*

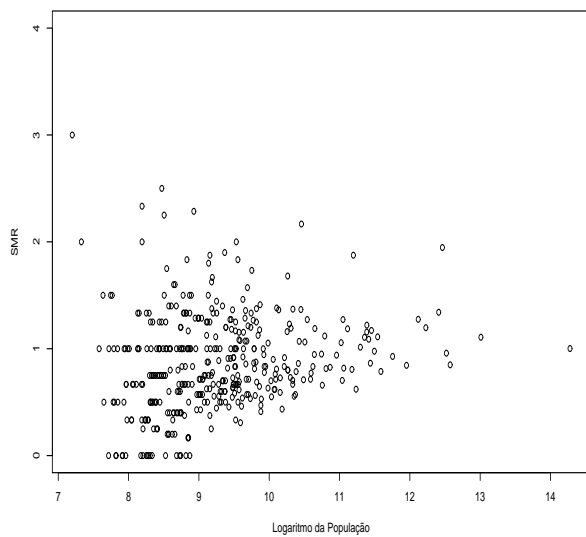


Figura 6.13: *Gráfico SMR versus logaritmo da população 2004*

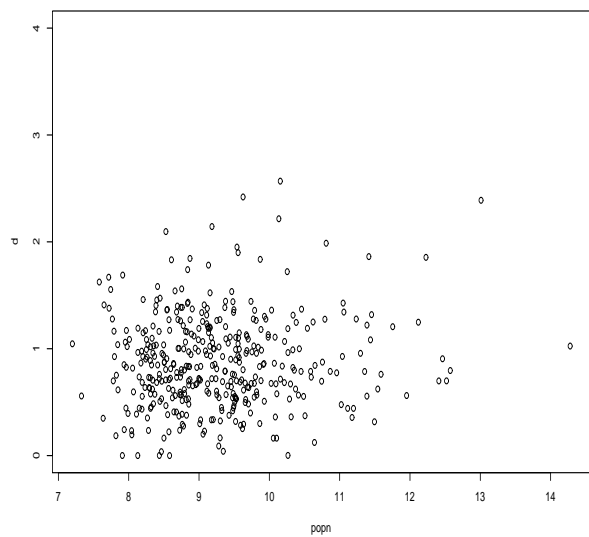


Figura 6.14: *Média a posteriori versus logaritmo da população em 2004*

A Figura 6.15 mostra um mapa com as médias a posteriori dos efeitos espaciais ( $\phi_{it}$ ) para  $t$  fixo igual a 2004. Percebemos que nos municípios de maior densidade populacional tendem a ter valores maiores para este parâmetro, o que vai implicar em valores maiores para o risco relativo. Estes efeitos espaciais visualizados no mapa indica a formação de regiões com maior grau de similaridade em relação a periculosidade. Áreas mais escuras do mapa mostram as regiões mais perigosas, em relação a situação de controle, e as áreas brancas mostram as menos perigosas.

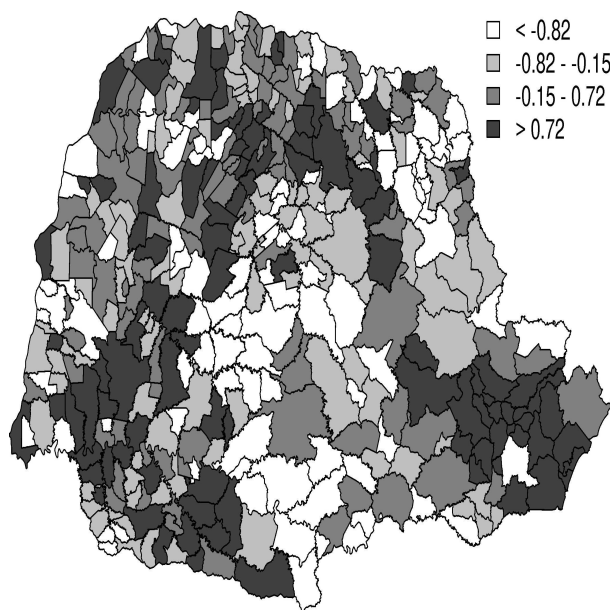


Figura 6.15: *Média a posteriori do componente estocástico estruturado do modelo selecionado em 2004*

Nas Figuras 6.16 e 6.17 apresentamos estimativas da média a posteriori do risco relativo do modelo selecionado, no mapa do estado do Paraná, para os anos de 1980 e 2004. Os municípios mais perigosos estão representados como áreas escuras no mapa. O critério utilizado para classificar um município como perigoso foi ele ter seu risco relativo significativamente superior a 1. No ano de 1980 o Estado do Paraná já apontava municípios como Foz do Iguaçu e São José dos Pinhais como problemáticos. A partir ano de 2004, aumenta a quantidade de municípios em estado crítico, ficando mais clara a formação de clusters em Londrina e região Metropolitana de Curitiba.

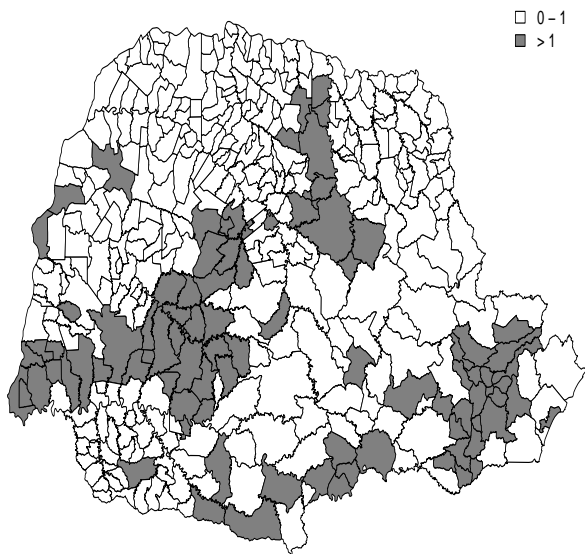


Figura 6.16: *Média a posteriori do risco relativo 1980 do modelo selecionado.*

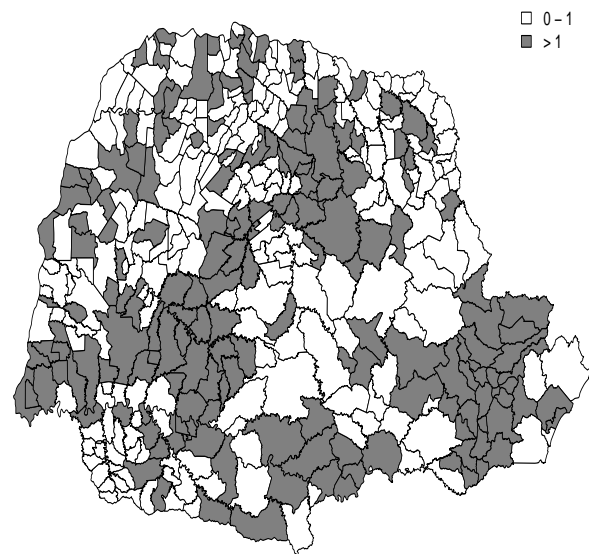


Figura 6.17: *Média a posteriori do risco relativo 2004 do modelo selecionado.*

No próximo capítulo serão propostos modelos multivariados espaciais com e sem estrutura espacial, com os resultados dos principais parâmetros do modelo selecionado pelo critério DIC. Dessa forma vamos poder visualizar se os modelos univariados e multivariados, detectam as mesmas regiões críticas.

# Capítulo 7

## Modelos Multivariados Espaciais

Dados multivariados são frequentemente observados em estudos de saúde ou segurança pública, pois em muitos casos variáveis como doenças, situação social e criminalidade são fortemente correlacionadas. Nessa seção propomos a utilização de modelos multivariados, tendo como variáveis resposta o número de óbitos por causas externas e o número de casos registrados de tráfico de drogas no estado do Paraná. Estudos de criminalidade apontam para uma alta correlação entre essas variáveis. Esta hipótese é sustentada pelo fato de se verificar o alto índice de mortes violentas próximo a locais de venda de drogas. Isto se deve ao fato do usuário de drogas muitas vezes cometer assaltos e até mesmo homicídios para poder consumir a droga. Com a análise multivariada podemos observar a evolução espacial da correlação e do risco relativo das variáveis de interesse. Além disso, o grau de periculosidade de cada região pode ser melhor definido analisando ambas as variáveis resposta. Como nos modelos univariados, vai ser assumido para a variável óbitos por causas externas uma distribuição Poisson. Assumimos também uma distribuição Poisson para a variável número de casos registrados de tráfico de drogas. Propomos o seguinte modelo:

$$y_{ik} \sim Poi(\Lambda_{ik})$$

$$\Lambda_{ik} = e_{ik}\psi_{ik},$$



onde o índice  $i$  identifica o município e  $k$  identifica a variável. Muitos trabalhos não levam em consideração a presença de correlação entre as variáveis e de maneira errônea assumem independência. Nesse trabalho vamos testar a hipótese de correlação nos componente estocásticos dessas duas variáveis, e caso seja constatado que a correlação não seja significativa, podemos propor duas modelagens independentes. Nesse capítulo analisamos apenas a estrutura espacial (e não a temporal), onde as variáveis serão estudadas no ano de 2004. O motivo para isso é que a coleta de dados da variável tráfico de drogas começou a partir do ano de 2003, e sendo assim não temos dados suficientes.

Vamos propor 2 modelos multivariados. Um deles sem estrutura espacial, e o outro estruturado no espaço. Esses modelos estão detalhados abaixo:

**Modelo 1:**  $\log(\psi_{ik}) = \alpha_k + \Phi_{ik}$ ,  $k = 1, 2$ , onde  $\alpha_k$  é o intercepto da variável  $k$ ,  $\Phi_i = (\Phi_{i,1}, \Phi_{i,2})^T$  é o componente estocástico bivariado. A priori vamos assumir que  $\Phi \sim CAR(\Lambda)$ . Note que essa distribuição é bivariada.  $\Lambda$  contém os parâmetros da variância, e assumimos para ela a priori uma distribuição Wishart. Para  $\alpha$  é necessário impor a priori uma distribuição Uniforme no intervalo  $(-\infty, \infty)$ , como comentado nos modelos univariados.

**Modelo 2:**  $\log(\psi_{ik}) = \alpha_k + \Phi_{ik}^*$ , onde  $\alpha_k$  é o intercepto da variável  $k$ , e vamos assumir que  $\Phi^* \sim N(0, \Lambda^*)$ . Note que esta é uma Normal Bivariada. Para  $\Lambda^*$  vamos assumir uma distribuição a priori Wishart. As distribuições Normal Multivariada e Wishart estão descritas no apêndice.

## 7.1 Inferência

A seguir apresentamos alguns resultados de inferência para os modelos 1 e 2 multivariados.

- Inferência sob o Modelo 1

Sejam  $\theta = (\alpha, \Phi, \Lambda)$ ,  $\alpha = (\alpha_1, \alpha_2)^T$ ,  $\Phi' = (\phi'_1, \phi'_2)$ , onde  $\phi'_1 = (\phi_{11}, \dots, \phi_{N1})$  e  $\phi'_2 = (\phi_{21}, \dots, \phi_{N2})$ .

Pelo Teorema de Bayes temos que:

$$p(\theta|y) \propto p(y|\theta)p(\theta),$$

$$p(\theta|y) \propto \prod_{i=1}^N \prod_{k=1}^2 \exp \{y_{ik} \log(e_{it}\psi_{ik}) - e_{ik}\psi_{ik} - \log(y_{ik}!)\} \prod_{k=1}^2 p(\alpha_k, \Phi, \Lambda)$$

$$p(\theta|y) \propto \prod_{i=1}^N \prod_{k=1}^2 \exp \{y_{ik} \log(e_{ik}\psi_{ik}) - e_{ik}\psi_{ik} - \log(y_{ik}!)\} \prod_{k=1}^2 p(\alpha_k)p(\Phi|\Lambda)p(\Lambda)$$

- Inferência sob o Modelo 2

Sejam  $\theta = (\alpha, \Phi^*, \Lambda)$ ,  $\alpha = (\alpha_1, \alpha_2)^T$ ,  $\Phi^* = (\phi'_1, \phi'_2)$ , onde  $\phi'_1 = (\phi_{11}, \dots, \phi_{N1})$  e  $\phi'_2 = (\phi_{21}, \dots, \phi_{N2})$

Pelo Teorema de Bayes temos que:

$$p(\theta|y) \propto p(y|\theta)p(\theta),$$

$$p(\theta|y) \propto \prod_{i=1}^N \prod_{k=1}^2 \exp \{y_{ik} \log(e_{it}\psi_{ik}) - e_{ik}\psi_{ik} - \log(y_{ik}!)\} \prod_{k=1}^2 p(\alpha_k, \Phi^*, \Lambda^*)$$

$$p(\theta|y) \propto \prod_{i=1}^N \prod_{k=1}^2 \exp \{y_{ik} \log(e_{ik}\psi_{ik}) - e_{ik}\psi_{ik} - \log(y_{ik}!)\} \prod_{k=1}^2 p(\alpha_k)p(\Phi^*|\Lambda^*)p(\Lambda^*)$$

## 7.2 Resultados Obtidos

Para realizar a inferência dos parâmetros foi utilizado o software WinBugs, pois a distribuição calculada na seção anterior não tem forma analítica fechada. Para o período de aquecimento da cadeia foram utilizadas 50000 iterações. Após essas iterações foram geradas mais 50000 iterações para fazer inferência. A convergência das cadeias foi observada através de uma inspeção visual. Pelo critério DIC obtivemos **3761.12** para o modelo 1 e **3769.13** para o modelo 2. Então o melhor modelo segundo esse critério foi o modelo 1, com componente estocástico estruturado. Os resultados mostrados a seguir foram obtidos sob esse modelo. Nas Figuras 7.1 e 7.2 foram construídos histogramas da média a posteriori dos interceptos para as variáveis morte por causas externas e número

de casos de tráfico, respectivamente. Nas Figuras 7.3 e 7.4 foram construídos histogramas da média a posteriori das variâncias dessas duas variáveis. Podemos observar que a estimativa da variância do número de casos de tráfico é bem maior que a estimativa da variância do número de mortes por causas externas, isto deve-se ao fato de que a média a posteriori do risco relativo da variável tráfico não varia de maneira tão suave como a média a posteriori da variável óbito por causa externa. Apresentamos na Figura 7.5 o histograma da média posteriori do parâmetro de correlação entre as duas variáveis resposta. Percebemos por essa figura que a correlação é significativa com média em torno de 0,4. Sendo assim, fica justificado modelar conjuntamente essas duas variáveis. Na Tabela 7.1 apresentamos a média e o desvio padrão a posteriori dos parâmetros de intercepto e variância para cada variável resposta e a correlação entre elas.

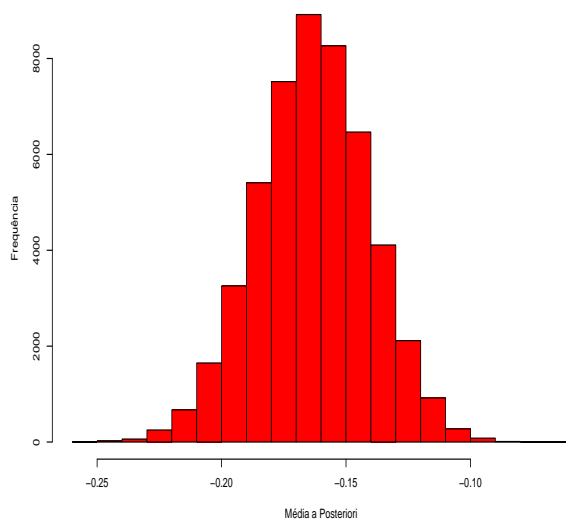


Figura 7.1: *Histograma da média a posteriori do intercepto da variável óbitos por causas externas*

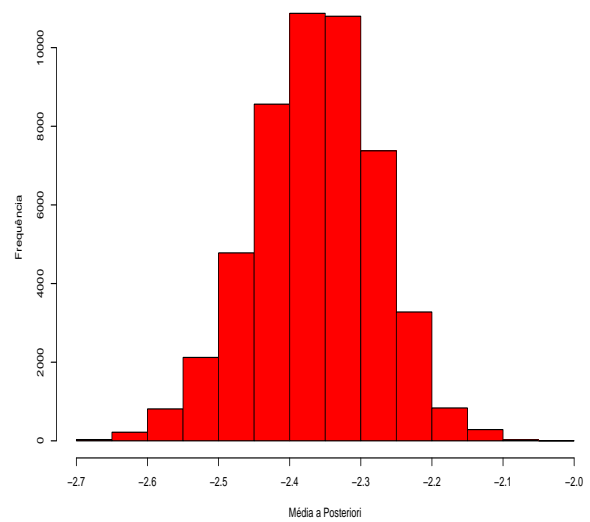


Figura 7.2: *Histograma da média a posteriori do intercepto da variável tráfico de drogas*

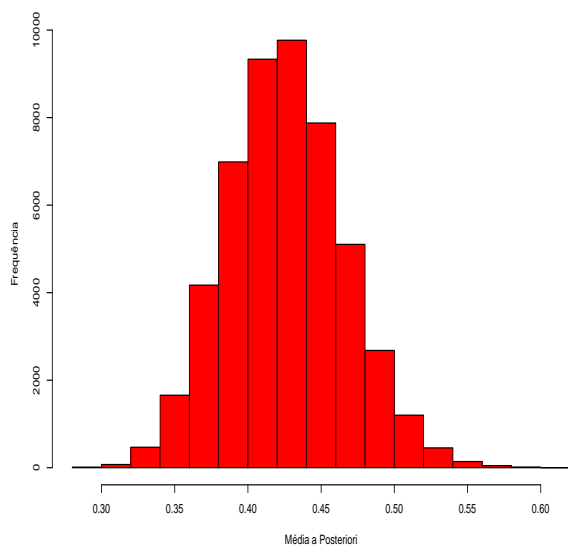


Figura 7.3: *Histograma da média a posteriori da variância da variável óbitos por causas externas.*

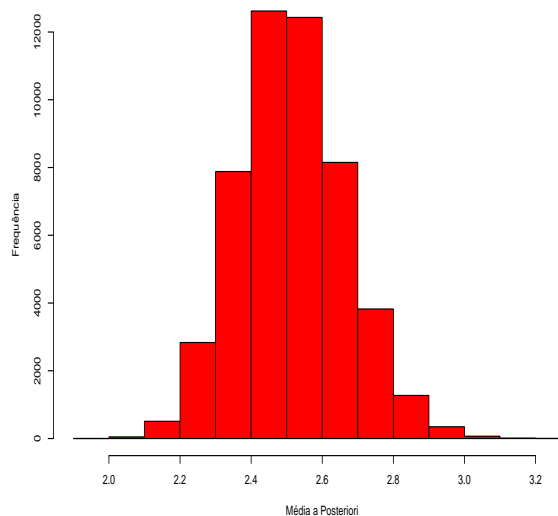


Figura 7.4: *Histograma da média a posteriori da variância da variável tráfico de drogas.*

Modelos	Média Posteriori	Desvio Padrão Posteriori
Intercepto Homicídio	-0.1641	0.02227
Intercepto Tráfico	-2.347	0.06784
Correlação	0.3771	0.08984
Variância Homicídio	0.4218	0.04465
Variância Tráfico	2.512	0.1567

Tabela 7.1: Resultados sob o Modelo 1

Nas Figuras 7.6 e 7.7 temos a média a posteriori do componente estocástico estruturado para o número de mortes por causas externas e tráfico, respectivamente. Podemos observar por essas figuras que o componente estocástico para o número de mortes por causas externas varia de forma suave no espaço, sendo o componente estocástico para o número de casos de tráfico mais volátil. Isto se deve ao fato de que a variância do componente estocástico da variável tráfico é maior que a variância do componente estocástico da variável óbito por causas externas.

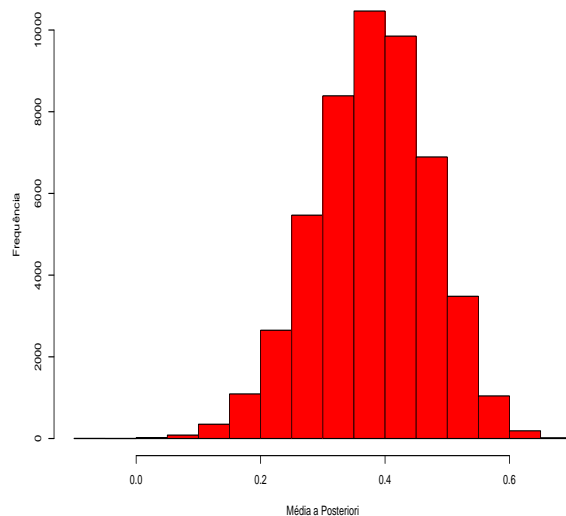


Figura 7.5: *Histograma da média a posteriori da correlação.*

Nas Figuras 7.8 e 7.9 Ficou clara a formação de clusters. A área pintada em branco no mapa indica os municípios com a melhor situação em relação a periculosidade no Paraná. As regiões com maiores riscos relativos, superiores a 1, concentram-se em Curitiba e sua Região Metropolitana e também em torno de Foz do Iguaçu e sua Região Metropolitana, que são os municípios mais violentos do estado. O município de Foz do Iguaçu faz fronteira com o Paraguai onde acontece um intenso comércio de tráfico de drogas, implicando em um elevado número de mortes por causas externas. Observamos uma correlação espacial da média a posteriori do risco relativo destas duas variáveis, comprovando a associação do tráfico com as mortes violentas.

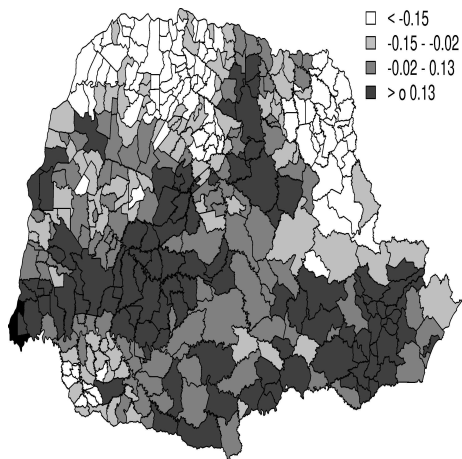


Figura 7.6: *Média a posteriori do componente estocástico estruturado de óbitos por causas externas.*

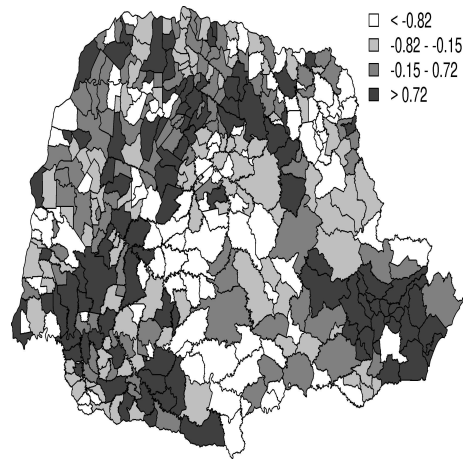


Figura 7.7: *Média a posteriori do componente estocástico estruturado do número de casos de tráfico de drogas.*

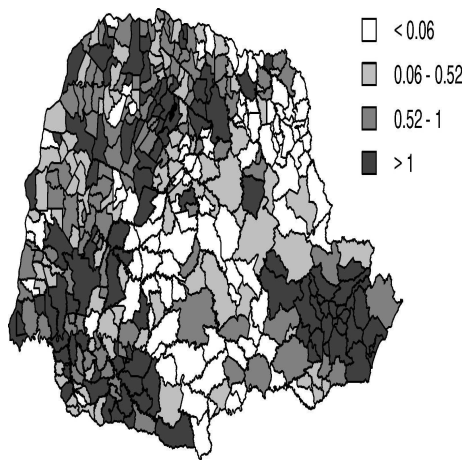


Figura 7.8: *Média a posteriori do risco relativo do número de casos de tráfico de drogas*

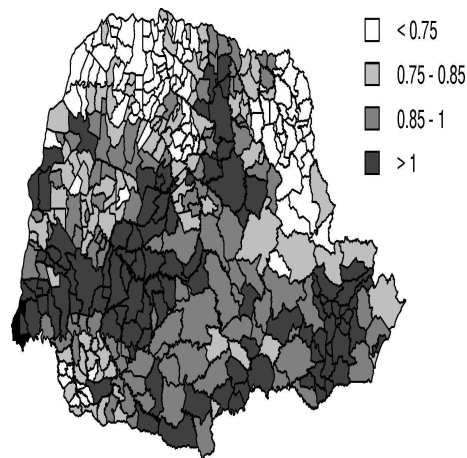


Figura 7.9: *Média a posteriori do risco relativo da variável óbitos por causas externas*

# Conclusão

Nesse trabalho propomos modelos univariados e multivariados para modelar o número de óbitos por causas externas no Estado do Paraná. Nos modelos univariados incorporamos variação espacial e temporal, e nos modelos multivariados apenas variação espacial, estando tempo fixo para o ano de 2004.

Concluimos através de análises descritivas que o fator populacional explica em grande parte a distribuição de óbitos no estado do Paraná. Através da comparação de diferentes modelagens univariadas, verificamos a necessidade de incorporar uma estrutura espacial para explicar o número de óbitos por causas externas.

Propusemos também uma modelagem bivariada para explicar o número de óbitos por causas externas e número de casos de tráfico de drogas, que são variáveis bastante correlacionadas. Ficou clara também a correlação espacial entre tráfico e óbitos por causas externas no estado do Paraná, ou seja, estas duas variáveis tem riscos relativos elevados em áreas geográficas similares, o que realmente justifica sua modelagem conjunta. Também nessa modelagem foi verificada a necessidade de trabalhar com componente estocástico estruturados no espaço. Na literatura, a utilização de modelos multivariados para modelar o componente estocástico através de uma estrutura CAR é recente. Como trabalho futuro poderíamos propor uma modelagem multivariada espaço-temporal.

Os modelos univariados e multivariados conseguiram detectar as regiões críticas do estado do Paraná em relação a violência. Como era de se esperar, os grandes centros urbanos, devido a problemas econômicos e sociais, despontaram como regiões problemáticas. Um ponto de interesse nesse trabalho foi corrigir superestimativas encontradas no SMR, onde municípios sem histórico de grande violência, foram apontados como regiões de risco devido a situações atípicas. Verificamos que esse problema foi corrigido pelos modelos

aqui propostos.

É importante ressaltar que as cidades com risco relativo inferior a 1 não são necessariamente exemplos de qualidade de segurança, porém queremos detectar quais as regiões que superam as expectativas de óbitos com base em sua população. Concluimos as regiões de Foz do Iguaçu e Região Metropolitana de Curitiba são os principais municípios com alto grau de periculosidade.



# Referências Bibliográficas

- [1] Anderson, B.D.O. e Moore, J.B. (1979). *Optimal filtering*. Englewood Cliffs, New Jersey: Prentice-Hall.
- [2] Assunção, R.M., Beato Filho, C.C., Silva, B.F.A., Marinho F.C. e Reis I.A. (2001). Conglomerados de homicídios e o tráfico de drogas em Belo Horizonte, Minas Gerais, de 1995 a 1999. *Caderno de Saúde Pública*, **17**, 1163-71.
- [3] Besag J.E., York, J.C. e Mollie, A. (1991). A Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**, 1-59.
- [4] Besag, J. e Kooperberg, C. (1995). On conditional and intrinsic autoregressions. *Biometrika*, **82**, 733-746.
- [5] Box, G. e Jenkins, G. (1976). *Time series Analysis: Forecasting and Control (segunda edição)*. San Francisco: Holden -Day.
- [6] Ferreira, G.S. e Schimdt, A.M. (2006). Spatial Modelling of the Relative Risk of Dengue Fever in Rio de Janeiro for the epidemic period between 2001 and 2002. *Brazilian Journal of Probability and Statistics*, Sao Paulo - Brasil, **20**, p. 29-47
- [7] Gamerman, D. and Lopes, H.F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. London: Chapman Hall.
- [8] Gamerman, D. e Moreira, A.R.B. e Rue, H. (2003). Space-varying regression models: specifications e simulation. *Computational Statistics Data Analysis*, **42**, 513-553.
- [9] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. Em J. Bernardo, J. Berger, A. Dawid e A. Smith (Editores), *Bayesian Statistics 4*, Oxford, 169-193, University Press.

- [10] Gelman, A. e Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, **7**, 457-511
- [11] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**, 721-741.
- [12] Gelfand, A.E. e Smith, A.M.F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**, 398-409
- [13] Harrison, P.J. e Stevens, C.F. (1976). Bayesian forecasting. *Journal of the Royal Statistical Society, B*, **38**, 205-247 (with discussion)
- [14] Harville, D.A. (1997). *Matrix Algebra from a Statistician's Perspective*. New York: Springer.
- [15] Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97-109.
- [16] Jeffreys, H. (1961). *Theory of Probability*. Oxford: Clarendon Press.
- [17] Kanh (2007). O fetiche das estatísticas criminais. *Fórum Brasileiro de Segurança Pública*.
- [18] Kypraios, T. (2003). Space-time Analysis of Leptospirosis Incidence in the City of Rio de Janeiro. Dissertação de Mestrado. Lancaster University, Inglaterra.
- [19] Ma, H. e Carlin, B.P. (2007). Bayesian multivariate areal wombling for multiple disease boundary analysis, *Bayesian Analysis*. **2**, Number 2, pp. 281-302
- [20] McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models*. Chapman and Hall: London.
- [21] 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.
- [22] Migon, H.S. e Gamerman, D. (1999). *Statistical Inference: an Integrated Approach*. London: Arnold.
- [23] Mollié, A. (1996). Bayesian Mapping of Disease. *Monte Carlo Markov Chain in Practice*. W.R. Gilks, S. Richardson and D.J. Spiegelhalter (editors).
- [24] Nelder, J.A. e Wedderburn, R.W.M. (1972). Generalized linear models. *Journal*

of the *Royal Statistical Society, A*, **135**,370-384.

[25] Nobre, A.A., Schimdt, A.M. e Lopes, H.F. (2005). Spatio-temporal models for mapping the incidence of malaria in Pará. *Environmetrics* (London), Reino Unido, v. 16, p. 291-304

[26] Paez, M.S. (2000). Estudo dos efeitos espaço-temporais na concentração de poluição na Região Metropolitana do Rio de Janeiro. Dissertação de mestrado, IM, Universidade Federal do Rio de Janeiro.

[27] Paez, M.S. (2004). Análise de modelos para a estimação e previsão de processos espaço-temporais. Tese de doutorado, IM, Universidade Federal do Rio de Janeiro.

[28] Paez, M.S. e Gamerman, D. (2003). Study of the space-time effects in the concentration of airborne pollutants in the Metropolitan Region of Rio de Janeiro. *Environmetrics*, **14**, 387-408.

[29] Paez, M.S. e Gamerman, D. (2005). *Modelagem de Processos espaço-temporais*. 1. ed. Vitória: Gráfica Universitária. 101 p.

[30] 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.

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

[32] Vivar-Rojas, J.C. (2004). Uma nova classe de modelos espaço-temporais para dados de área. Dissertação de mestrado, IM-UFRJ.

[33] West, M., Harrison, P.J. e Migon, H.S. (1985). Dynamic generalized linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association*, **80**, 73-97.

[34] West, M. e Harrison, P.J. (1997). *Bayesian Forecasting and Dynamic Models (segunda edição)*. Springer Verlag, New York.

# Apêndice

No Apêndice será exibido os comandos para construção dos Mapas no software R, a construção dos intervalos de máxima densidade a posteriori, utilizando o pacote Coda também no software R. Também será mostrado o código para inferência dos modelos no WinBugs e no final do Apêndice estão descritas as distribuições Normal Multivariada e Wishart.

## Comandos R e WinBugs

### Comandos Mapas

```
require(maptools)
dados=read.table(arquivo.txt",head=T,dec=",")
attach(dados)
mappolys = Map2poly(map.pr,as.character(map.pratt.dataCODE))
perc = round(quantile(V1, probs=seq(0,1,0.25)), digits=2)
colours= sort(gray(0:4 / 4),decreasing=T)
plot(mappolys, col=colours[findInterval(V1, perc)],axes=F)
legend(x=c( -49,-48.7),y=c( -22.5, -22.3),legend=leglabs(perc),fill=colours,bty="n")
dev.off( )
```

### Comandos Intervalo Máxima Densidade a Posteriori

```
require(coda)
dados=read.coda("beta.txt", "ind.txt")

ic <- HPDinterval(dados)
plot(1:nrow(ic), ylim=range(ic),xlab="tempo",ylab="media a posteriori do coeficiente",
```

```
type="n")
sapply(1:nrow(ic), function(i)
arrows(i, ic[i,1], i, ic[i,2], code=3, angle=70))
```

## Comandos Winbugs Modelo Multivariado

```
model {  
S[1:2, 1:399] mv.car(adj[], weights[], num[], omega[ , ])  
for (i in 1:399) {  
for (k in 1:2) {  
y[k,i] dpois(mu[k,i])  
log(mu[k,i]) <- log(esperado[k,i]) + alpha[k] + S[k, i]  
}  
RR1[i] <- exp(alpha[1] + S[1,i])  
RR2[i] <- exp(alpha[2] + S[2,i])  
}  
for (k in 1:2) {  
alpha[k] < dflat()  
}  
R[1,1]<-0.1  
R[1,2]<-0  
R[2,1]<-0  
R[2,2]<-0.1  
omega[1:2, 1:2] dwish(R[ , ], 2)  
sigma2[1:2, 1:2] <- inverse(omega[ , ])  
sigma[1] <- sqrt(sigma2[1, 1])  
sigma[2] <- sqrt(sigma2[2, 2])  
corr <- sigma2[1, 2] / (sigma[1] * sigma[2])  
}
```

## Distribuições Multivariadas

1. Distribuição Whishart  $f_W(w) = \frac{|W|^{\frac{n-p-1}{2}}}{2^{\frac{np}{2}} |V|^{\frac{n}{2}} \Gamma_p(\frac{n}{2})} \exp(\frac{-1}{2} Tr(V^{-1}W))$

onde  $W$  é uma matriz  $p \times p$  positiva definida,  $V$  é uma matriz  $p \times p$  de escala positiva definida e  $n$  são os graus de liberdade.

2. Distribuição Normal Multivariada

$$f_X(x_1, \dots, x_N) = \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma|^{\frac{1}{2}}} \exp(\frac{-1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu))$$

onde,  $\mu = (\mu_1, \dots, \mu_n)$  é o vetor de médias,  $\sigma$  é a matriz de covariância (positiva definida e com dimensão  $N \times N$ ).

# 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](#)