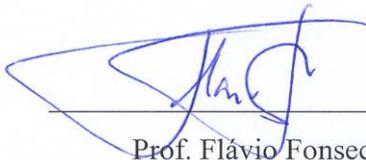


MODELAGEM DOS ACIDENTES DE TRÂNSITO COM VÍTIMAS
NO MUNICÍPIO DO RIO DE JANEIRO

Cosme Marcelo Furtado Passos da Silva

TESE SUBMETIDA AO CORPO DOCENTE DA COORDENAÇÃO DOS
PROGRAMAS DE PÓS-GRADUAÇÃO DE ENGENHARIA DA UNIVERSIDADE
FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS
NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR EM CIÊNCIAS EM
ENGENHARIA BIOMÉDICA.

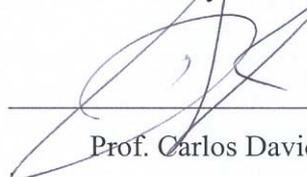
Aprovada por:



Prof. Flávio Fonseca Nobre, Ph.D



Profa. Rosimary Terezinha de Almeida, Ph.D



Prof. Carlos David Nassi, Dr.-Ing.



Profa. Rejane Sobrino Pinheiro, D.Sc.



Profa. Maria Cecília de Souza Minayo, D.Sc.

RIO DE JANEIRO, RJ - BRASIL

MARÇO DE 2006

SILVA, COSME MARCELO FURTADO PASSOS DA

Modelagem dos Acidentes de Trânsito com Vítimas
no Município do Rio de Janeiro [Rio de Janeiro] 2006

XIV, 133 p. 29,7 cm (COPPE/UFRJ, D.Sc.,
Engenharia Biomédica, 2006)

Tese - Universidade Federal do Rio de Janeiro,
COPPE

1. Acidentes de Trânsito
2. Modelo de Regressão Binomial Negativa
3. Análise Espacial

I. COPPE/UFRJ II. Título (série)

Ao meu pai **Maciel** Pedro Passos Soares da Silva (*in memorian*) e a minha mãe **Marlene** Furtado da Silva, por sempre me apoiarem e dedicarem todo o seu amor. Eu amo vocês.

À minha irmã Maria José Passos Furtado de Avellar (**Zezé**).

Ao meu sobrinho e afilhado **Victor Hugo** Passos Furtado de Avellar.

AGRADECIMENTOS

A Deus, pois sem a sua Luz nada seria possível.

Ao Prof. Dr. Flávio Fonseca Nobre pela sua orientação, por tudo que conseguimos realizar nesse trabalho e pelas exigências que contribuíram não apenas para a melhora da qualidade do trabalho, mas principalmente para a minha formação.

À Profa. Dra. Rosimary Terezinha de Almeida pelas sugestões, discussões e incentivo e também por participar desta etapa decisiva.

Ao Prof. Dr. Carlos David Nassi, por ter cedido as variáveis relacionadas ao trânsito e por participar da banca examinadora.

À Profa. Dra. Rejane Sobrino Pinheiro por participar da banca examinadora e pelas sugestões.

À Profa. Dra. Maria Cecília de Souza Minayo, por ter me apresentado aos estudos da violência e seu impacto na área da Saúde, por sempre me incentivar e apoiar e por participar da banca examinadora.

À Profa. Dra. Ana Beatriz Monteiro Fonseca pelas sugestões, pela revisão nos conceitos estatísticos, e por participar da banca examinadora do exame de qualificação.

Aos meus pais, Maciel Pedro Passos Soares da Silva (*in memorian*) e Marlene Furtado da Silva, a minha irmã Maria José Passos Furtado de Avellar (Zezé) e ao meu sobrinho Victor Hugo Passos Furtado de Avellar, pela generosidade e presença constantes.

Aos grandes amigos do Centro Latino Americano de Estudos de Violência e Saúde Jorge Careli (CLAVES/ENSP/FIOCRUZ), que sempre me apoiaram e incentivaram e me mostraram a importância de estudos relacionados aos eventos violentos.

À Profa. Dra. Denise Britz do Nascimento e Silva pela leitura e contribuições no texto das bases teóricas.

A todos os amigos da Coordenação de Serviços e Comércio (COSEC/DPE/IBGE), e em especial à minha ex-chefe e eterna amiga Maria Deolinda Borges Cabral, pelo apoio sem o qual a conclusão deste trabalho certamente seria mais difícil.

A todos os amigos da Escola Nacional de Ciências Estatísticas (ENCE/IBGE).

Ao Chefe de Polícia Civil do Estado do Rio de Janeiro, Dr. Álvaro Lins dos Santos por ter fornecido os dados.

À Profa. Andréa Diniz da Silva, pela ajuda na obtenção das informações do Censo Demográfico.

Ao Prof. Aurélio Ferreira Braga, pela ajuda na obtenção das variáveis relacionados ao trânsito.

A Profa. Dra. Simone Gonçalves de Assis e a Profa. Maria Fernanda Pimentel de Maria da Silva pela ajuda na elaboração do resumo em inglês.

Aos amigos da COPPE, professores, funcionários e alunos, pela colaboração constante.

Ao meu amigo Saint-Clair dos Santos Gomes Júnior por suas sugestões e discussões sobre estatística.

À minha grande amiga Liana Wernersbach Pinto, por toda a sua amizade e ajuda nas buscas de referências bibliográficas, nas discussões e no ouvido amigo para escutar meus lamentos.

A todos os meus amigos que puderam compreender o meu afastamento nesse período.

A todos aqueles que contribuíram direta ou indiretamente para a realização deste trabalho, expesso meus sinceros agradecimentos.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

MODELAGEM DOS ACIDENTES DE TRÂNSITO COM VÍTIMAS

NO MUNICÍPIO DO RIO DE JANEIRO

Cosme Marcelo Furtado Passos da Silva

Março/2006

Orientador: Flávio Fonseca Nobre

Programa: Engenharia Biomédica

Este trabalho teve como principal objetivo realizar um estudo dos acidentes de trânsito com vítimas no Município do Rio de Janeiro, visando identificar o seu contexto socioespacial, utilizando os Registros de Ocorrência, da Secretaria de Estado de Polícia Civil do Rio de Janeiro, no período de janeiro de 1998 a dezembro de 2000. Foram utilizadas técnicas de análise estatística espacial para caracterizar a distribuição espacial dos acidentes, considerando como unidades geográficas os bairros que compõem o município. Os modelos de regressão Binomial Negativa foram utilizados para descrever o número de acidentes de trânsito com vítimas, o número de atropelamentos e o número de colisões com vítimas tendo como variáveis explicativas características ambientais, socioeconômicas e demográficas. Este trabalho conseguiu identificar as principais variáveis relacionadas à ocorrência desses acidentes e pode servir como base para a criação de um instrumento de apoio à decisão para o controle e redução desses eventos.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

TRAFFIC ACCIDENTS WITH VITIMS MODELLING
IN RIO DE JANEIRO CITY

Cosme Marcelo Furtado Passos da Silva

March/2006

Advisor: Flávio Fonseca Nobre

Department: Biomedical Engineering

This work aims to study traffic accidents with victims in the city of Rio de Janeiro, describing its social-spatial context. It was used the incident registers of the Civil Police State Secretary of Rio de Janeiro, from January of 1998 to December of 2000. Techniques of Spatial Statistic analysis were used to characterize the spatial distribution of the accidents, considering as geographic unities the county districts. The Negative Binomial Regression models were used to describe the number of traffic accidents with victims, the number of road accidents and the number of collisions with victims having as explanatory variables, social, economical and demographic characteristics. This work succeeded in identifying the main variables related to these accidents and it can be useful to create an instrument to support the decision for the control and reduction of these events.

ÍNDICE

1. Introdução e Justificativa	1
2. Revisão da Literatura	6
3. Bases Teóricas	20
3.1. Análise Espacial	20
3.2. Modelos Lineares Generalizados	26
3.2.1. Modelos de Regressão de Poisson	42
3.2.2. Modelos de Regressão Binomial Negativa	44
4. Material e Métodos	46
5. Resultados	53
5.1. Análise Exploratória	53
5.2. Análise Espacial	57
5.2.1. Análise Espacial dos Acidentes de Trânsito com Vítimas	58
5.2.2. Análise Espacial dos Atropelamentos	64
5.2.3. Análise Espacial das Colisões com Vítimas	68
5.3. Modelagem dos Dados	75
5.3.1. Acidentes de Trânsito com Vítimas	82
5.3.2. Atropelamentos	88
5.3.3. Colisões com Vítimas	92

6. Discussão	97
7. Conclusões	107
8. Referências Bibliográficas	110

Anexos

Anexo 1: Município do Rio de Janeiro e seus bairros	121
Anexo 2: Distribuição do total de acidentes de trânsito com vítimas do Município do Rio de Janeiro, no período de 1998 a 2000, segundo bairro e ano de ocorrência.	122
Anexo 3: Distribuição do total de acidentes de trânsito com vítimas do Município do Rio de Janeiro, no período de 1998 a 2000, segundo bairro de ocorrência e tipo de evento.	128

ÍNDICE DE FIGURAS

1. Diagrama de espalhamento de Moran.	25
2. Distribuição dos registros segundo ano de notificação.	53
3. Distribuição dos registros segundo tipo de evento.	54
4. Distribuição dos registros segundo dia da semana.	55
5. Distribuição dos registros segundo hora do evento.	56
6. Boxplot para a distribuição do número de ocorrências de acidentes de trânsito com vítimas, de colisões com vítimas e de atropelamentos.	58
7. Distribuição por quintis dos acidentes de trânsito com vítimas. Triênio 1998-2000.	59
8. Distribuição por intervalos dos acidentes de trânsito com vítimas. Triênio 1998-2000.	60
9. Médias móveis por quintis dos acidentes de trânsito com vítimas. Triênio 1998-2000.	61
10. Médias móveis por intervalos dos acidentes de trânsito com vítimas. Triênio 1998-2000.	62
11. Diagrama de espalhamento de Moran para os acidentes de trânsito com vítimas. Triênio 1998-2000.	63
12. Distribuição por quintis dos atropelamentos. Triênio 1998-2000.	64
13. Distribuição por intervalos dos atropelamentos. Triênio 1998-2000.	65

14. Médias móveis por quintis dos atropelamentos. Triênio 1998-2000.	66
15. Médias móveis por intervalos dos atropelamentos. Triênio 1998-2000.	66
16. Diagrama de espalhamento de Moran para os atropelamentos. Triênio 1998-2000.	68
17. Distribuição por quintis das colisões com vítimas. Triênio 1998-2000.	69
18. Distribuição por intervalos das colisões com vítimas. Triênio 1998-2000.	70
19. Médias móveis por quintis das colisões com vítimas. Triênio 1998-2000.	71
20. Médias móveis por intervalos das colisões com vítimas. Triênio 1998-2000.	72
21. Diagrama de espalhamento de Moran para as colisões com vítimas. Triênio 1998-2000.	73
22. Boxplot para a distribuição da variável pop dos bairros do Município do Rio de Janeiro, em 2000.	76
23. Boxplot para a distribuição da variável area dos bairros do Município do Rio de Janeiro.	76
24. Boxplot para a distribuição das variáveis idh, sanea, menor18, maior60, homem, casado e superior dos bairros do Município do Rio de Janeiro, em 2000.	77
25. Boxplot para a distribuição da variável renda dos bairros do Município do Rio de Janeiro, em 2000.	77
26. Boxplot para a distribuição da variável velocmed dos bairros do Município do Rio de Janeiro.	78

27. Boxplot para a distribuição das variáveis via_ap , via_as , via_c e via_e dos bairros do Município do Rio de Janeiro.	78
28. Boxplot para a distribuição da variável via_l dos bairros do Município do Rio de Janeiro.	79
29. Boxplot para a distribuição da variável lin_bus dos bairros do Município do Rio de Janeiro.	79
30. Boxplot para a distribuição da variável $intersec$ dos bairros do Município do Rio de Janeiro.	80
31. Análise de resíduos para o modelo dos acidentes de trânsito com vítimas com interações $via_c*pop+lin_bus+via_as+menor18$.	86
32. Análise de resíduos para o modelo dos acidentes de trânsito com vítimas com interações $via_c+pop*lin_bus+via_as+menor18$.	87
33. Análise de resíduos para o modelo dos atropelamentos com interações $pop*lin_bus+via_as+menor18$.	91
34. Análise de resíduos para o modelo das colisões com vítimas com interações $via_c*via_as+lin_bus+menor18$.	95

ÍNDICE DE QUADROS

1. Variáveis respostas e distribuição pertencente à família exponencial na forma canônica do componente aleatório dos modelos lineares generalizados.	30
2. Identificadores da família exponencial das distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.	30
3. Funções de ligação canônicas das distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.	32
4. Função desvio para as distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.	36
5. Principais tipos de resíduos utilizados em Modelos Lineares Generalizados	40
6. Estatísticas descritivas (mínimo, máximo, quartil 1, mediana, quartil 3, média e desvio padrão) das variáveis explicativas utilizadas na modelagem.	75
7. Estatísticas descritivas (mínimo, máximo, média e variância) das variáveis acidentes de trânsito com vítimas, colisões com vítimas e atropelamentos.	82
8. Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo para acidentes de trânsito com vítimas.	86
9. Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo para atropelamentos.	90

10. Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo par a colisões com vítimas. 95

1. INTRODUÇÃO E JUSTIFICATIVA

Na última década, os eventos violentos foram responsáveis pelas principais mudanças ocorridas no quadro de morbi-mortalidade do país. A partir da década de 80, as mortes por causas externas aumentaram em todas as faixas etárias atingindo, principalmente, a parcela economicamente ativa da população de todas as regiões metropolitanas (SOUZA e MINAYO, 1994), destacando-se como um dos principais problemas na área da saúde. Esse tipo de mortalidade que no início da década de 80 ocupava o quinto lugar entre as causas de mortalidade, passou, no final da década, a ocupar o segundo lugar como causa de morte (MINAYO, 1994). Desde 1989, as causas externas constituem a segunda causa de morte entre os brasileiros e a primeira causa de óbito para aqueles entre 5 e 39 anos (MELLO JORGE e LATORRE, 1994). Em 1998, esse grupo de causas representou 68% do total de óbitos dos jovens brasileiros com idades entre 15 e 19 anos, chegando a 76,7% no sexo masculino para a mesma faixa etária (SIMÕES, 2001).

Devido à complexidade que envolve os múltiplos eventos designados pela categoria violência na área da Saúde Pública (ROSENBERG *et al.*, 1992), eles recebem na décima revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (OMS, 1996), a denominação de “causas externas” (V01 – Y98). Esta definição abrange uma gama extremamente extensa de fenômenos: desde todos os tipos de acidentes (domésticos, transporte, trânsito e de trabalho) até homicídios, suicídios, e mesmo as lesões resultantes de operações de guerra e ação policial. A falta de especificidade dificulta o aprofundamento sobre questões tão distintas, sendo

necessário um detalhamento por causas específicas, que passam então a constituir um referencial mais estratégico para o debate na Saúde Pública.

Embora os homicídios sejam a principal causa externa específica, com taxas de mortalidade por 100.000 habitantes variando de 5,16 óbitos no Piauí a 58,77 óbitos em Pernambuco, em 1998 (NUNES *et al.*, 2001), os acidentes de trânsito contribuem em muito para o incremento no conjunto de óbitos por causas violentas. No Estado do Rio de Janeiro, em 1980, os homicídios representaram 33,4% dos óbitos por violência e os acidentes de trânsito 30,4%, enquanto em 1988 foram, respectivamente, 45,2% e 31,2% (MINAYO, 1994).

As taxas de mortalidade por acidentes de trânsito, no município do Rio de Janeiro, são bem mais elevadas que as do Brasil (MELLO JORGE e LATORRE, 1994). Para 1987, na faixa de 0 a 4 anos, as taxas (por 100.000 habitantes) eram de 6,5 no Rio de Janeiro e 4,8 no Brasil. Na faixa dos 5 aos 9 anos, esta diferença é ainda mais intensa: 14,2 e 7,6, respectivamente. Segundo MINAYO (1994), a violência no trânsito é o ponto negro mais destacado nas estatísticas de causas externas no Brasil. Essa autora, chama a atenção para a incidência sobre a faixa de 5 a 14 anos, idade escolar, onde é a primeira causa de morte.

Mesmo com a implantação do Código de Trânsito Brasileiro (BRASIL, 1997), em 22 de janeiro de 1998, a partir da Lei nº 9.503, sancionada em 23 de setembro de 1997 e publicada no Diário Oficial da União em 24 de setembro de 1997, as estradas brasileiras matam de 10 a 70 vezes mais que as rodovias do grupo dos sete países mais ricos do mundo composto por Alemanha, Canadá, Estados Unidos, França, Grã-Bretanha, Itália e Japão, e denominado G-7. A cada trecho de mil quilômetros, morrem

213 pessoas nas estradas brasileiras, enquanto a média para o G-7 é de 10,7 óbitos anuais, e particularmente no Canadá é de 3 óbitos anuais para cada trecho de mil quilômetros (ADURA, 2002).

Diante do exposto, observa-se a necessidade de estudos mais amplos e consistentes acerca dos óbitos por acidentes de trânsito e também sobre a morbidade, visto que esses eventos vêm se mostrando muito freqüentes nos últimos anos, tornando-se, assim, um grande problema para a Saúde Pública, e porque poucos trabalhos têm buscado identificar as suas condições facilitadoras.

Torna-se necessário, então, tentar detectar quais as áreas mais freqüentemente atingidas, e também suas características socioeconômicas e demográficas, para que a partir dessas informações sejam tomadas medidas que visem a prevenção dos acidentes de trânsito.

Entre as medidas já implementadas em outros países que obtiveram respostas satisfatórias e outras, apontadas na literatura por GRISCI (1991), PORTUGAL e SANTOS (1991), LEITÃO (1997), BASTOS *et al.* (1999), HIJAR-MEDINA *et al.* (1999) e MARIN e QUEIROZ (2000), destacam-se:

- (1) realização de campanhas permanentes de segurança no trânsito com participação comunitária, no que tange ao excesso de velocidade e à ingestão de bebidas alcoólicas;
- (2) melhorias da capacitação dos diversos condutores de veículos, possibilitando uma “formação continuada” aos motoristas, especialmente dos motoristas profissionais;
- (3) aumento da segurança dos veículos;

- (4) melhorias nas sinalizações das vias, além de melhor policiamento e fiscalização, principalmente nos pontos de maior risco;
- (5) maior controle sobre as cargas de trabalho dos motoristas profissionais;
- (6) educação do pedestre quanto aos seus direitos e deveres, pois a eles também se aplicam as leis de trânsito;
- (7) participação da população e comunidades na elaboração das políticas, campanhas e planejamento do trânsito.

Na literatura, a maioria dos estudos analíticos sobre acidentes de trânsito utiliza a construção de modelos buscando identificar as variáveis explicativas para o número de acidentes de trânsito (ABDEL-ATY e RADWAN, 1999; FRIDSTROM *et al.*, 1995; GREIBE, 2003; JEGEDE, 1988; MIAOU, 1994; MIAOU e LUM, 1993; SHANKAR *et al.*, 1995), para as taxas de acidentes de trânsito (LASCALA, 2000), e/ou para as taxas de mortalidade por acidentes de trânsito (KMET *et al.*, 2003).

Entre os modelos mais utilizados para descrever os acidentes de trânsito pode-se destacar os Modelos de Regressão Linear Múltiplos (LASCALA *et al.*, 2000; JEGEDE, 1988; KMET *et al.*, 2003; MIAOU e LUM, 1993), Modelos de Regressão de Poisson (MIAOU, 1994; MIAOU e LUM, 1993; SHANKAR *et al.*, 1995; WOOD, 2005), Modelos de Regressão de Poisson inflacionado de zeros (MIAOU, 1994) e os Modelos de Regressão Binomial Negativa (ABDEL-ATY e RADWAN, 1999; WOOD, 2005; MIAOU, 1994; SHANKAR *et al.*, 1995; CHANG, 2005).

Sendo assim, este trabalho tem como principal objetivo realizar um estudo dos acidentes de trânsito com vítimas no Município do Rio de Janeiro, visando identificar o seu contexto socioespacial.

Esta proposta inclui a utilização de técnicas de análise estatística espacial para caracterizar a distribuição espacial dos acidentes de trânsito com vítimas no Município do Rio de Janeiro, considerando como unidades geográficas os bairros que compõem o município.

Além disso, através da construção de modelos que descrevam o número de acidentes de trânsito com vítimas utilizando como variáveis explicativas as características ambientais, socioeconômicas e demográficas, pretende-se identificar as principais variáveis facilitadoras da ocorrência desses agravos, buscando assim, um ferramental de apoio à decisão para o controle e redução dessas ocorrências.

2. REVISÃO DA LITERATURA

A ocorrência de acidentes de trânsito vem se tornando, ao longo dos anos, um grande problema na área da Saúde Pública. Esse grupo específico ocupava a primeira posição no conjunto de mortes por causas externas no Brasil, porém a partir dos anos 90, passou a ocupar a segunda posição, cedendo lugar para os homicídios (MELLO JORGE *et al.*, 2001).

No conjunto das causas externas, os homicídios foram a principal causa de mortes masculinas com uma taxa de 44,82 óbitos por 100.000 homens, no Brasil, em 1998, enquanto no sexo feminino os acidentes de transportes, que são todos os acidentes que envolvem veículos destinados ao transporte de pessoas ou mercadorias de um lugar para outro, foram a principal causa externa específica, responsáveis por uma taxa de 7,3 óbitos por 100.000 mulheres (SOUZA *et al.*, 2003).

De acordo com os dados de mortalidade do país (CENEPI, 1997), em 1994, dos 887.594 óbitos, 107.292 (12,1%) foram devido às causas externas e, destes, 29.014 corresponderam a acidentes de trânsito (3,3% do total de mortes e 27,0% das causas externas). Dos 116.942 óbitos ocorridos no estado do Rio de Janeiro, em 1994, 17.663 (15,1%) foram devido às causas externas, sendo que 2.213 foram devido aos acidentes de trânsito (12,5% das causas externas e 1,9% da mortalidade geral do estado).

No período de 1979 a 1995, em Recife, observou-se que no conjunto das causas externas, os acidentes de trânsito tiveram um comportamento mais estável que as outras causas específicas. Contudo, no grupo etário de 0 a 19 anos verificou-se um aumento de

28,9% no coeficiente de mortalidade dos acidentes de trânsito para o mesmo período (BARROS, 2001).

No conjunto das causas externas, os acidentes de trânsito e transportes foram as principais causas de mortes de pessoas com 60 ou mais anos de vida, no Brasil, nos anos de 1980 e 1998. Esse mesmo comportamento foi verificado para as capitais de regiões metropolitanas de Belém, Fortaleza, Recife, Belo Horizonte, Rio de Janeiro e São Paulo. Em Curitiba e Porto Alegre essa causa específica foi a principal somente em 1998, uma vez que em 1980 as causas de intenção indeterminada apresentaram maiores proporções. Em Salvador as causas não especificadas foram as principais causas de mortes de idosos, no conjunto das causas externas, em 1980 e 1998, seguidas dos acidentes de trânsito e transportes (SILVA, 2000). Deve-se ressaltar que a mortalidade por causas externas entre os idosos apresenta um comportamento diferente da mortalidade por causas externas nos jovens. Entre os idosos as principais causas específicas são os acidentes de trânsito e transportes e as quedas, enquanto nos jovens os homicídios são as principais causas (MINAYO, 1994, PAIM et al., 1999).

Estudos como os de MELLO JORGE (1990), SOUZA (1992 e 1993) e LAIT (1992), que analisaram detidamente as falhas no preenchimento da causa básica do óbito pelo IML, sugerem que a maioria dos óbitos classificados no grupo das lesões ignoradas, na verdade, são mortes por acidentes de trânsito e homicídios não esclarecidas. SOUZA *et al.* (1997) relatam existir uma grave falha quanto ao esclarecimento da causa básica do óbito no Município do Rio de Janeiro. Os autores chamam a atenção para o fato de 47,1% das mortes por causas externas, entre 1980 e 1994, na faixa de 10 a 19 anos, no Município do Rio de Janeiro, terem sido incluídas no grupo das lesões ignoradas, indicando que as taxas de mortalidade por homicídios,

assim como por acidentes de trânsito poderiam ser ainda maiores caso houvesse um preenchimento mais criterioso das declarações de óbito.

Com o objetivo de avaliar a concordância das informações sobre causa básica de óbito por acidentes de trânsito de cinco hospitais de Belo Horizonte presentes nas declarações de óbitos e comparando-as com as codificações feitas a partir das informações obtidas nos mesmos hospitais quando os pacientes foram admitidos, LADEIRA e GUIMARÃES (1998), verificaram entre os 50 óbitos ocorridos no período de 3 de novembro a 3 de dezembro de 1994, para os quais foi possível obter as duas informações, uma baixa concordância. Os autores destacam um elevado percentual (32%) de óbitos classificados nas declarações de óbitos como “Outras causas não especificadas, ambientais e naturais”, ocasionando assim uma grande subnotificação no conjunto dos acidentes de trânsito. Outro fato que chama a atenção é que 38% dos óbitos foram classificados como “Acidentes de veículos a motor não especificados, vítima não especificada”, não trazendo qualquer esclarecimento sobre o tipo de acidente e da vítima. Destaca-se ainda, um aumento percentual nos “Atropelamentos” que passaram de 24% para 46% dos óbitos e das “Colisões entre veículos a motor” que subiram de 2% para 24% dos óbitos.

Em 1995, as causas externas foram a segunda causa de morte da população carioca (15,1% dentre todas as causas de morte) e a principal causa de morte prematura que é aquela que ocorre numa etapa da vida de altas criatividade e produtividade anterior a esperança de vida da população (REICHENHEIM e WERNECK, 1994; SCHRAMM *et al.*, 2004). Neste mesmo ano, o município registrou 1.064 mortes por acidentes de trânsito, o que significou o segundo lugar das mortes no grupo das causas externas (excluindo-se as “causas ignoradas se acidental ou intencionalmente

infligidas”). Cabe ressaltar que, neste ano, a Secretaria Municipal de Saúde melhorou a qualidade do preenchimento dos atestados de óbito por causas externas. Como resultado, o percentual dos óbitos em que se ignorava a causa ou origem da lesão foi reduzido de 40% para cerca de 10% (SMS-RJ, 1996).

Considerando-se os anos potenciais de vida perdidos (APVP - indicador que qualifica a ocorrência precoce de mortes em relação a uma expectativa de vida de 70 anos), em 1990, as causas externas foram responsáveis por 39,2% das mortes prematuras no estado e 36,8% no município do Rio de Janeiro. Os acidentes de trânsito representaram, neste ano, respectivamente, para o município e o estado do Rio de Janeiro, 8,3% (terceira causa) e 7,6% (quarta causa) dos APVP (REICHENHEIM e WERNECK, 1994).

Embora não se disponha de muitos dados de morbidade por acidentes de trânsito, sabe-se que estes eventos constituem parcela significativa dos incidentes violentos ocorridos no país, sobretudo nas capitais. Segundo ASSIS e SOUZA (1995), ao efetuarem um levantamento de todas as ocorrências policiais não fatais que envolveram crianças e adolescentes do Município do Rio de Janeiro, no ano de 1990, verificaram neste período, que os acidentes de trânsito e transportes foram os eventos que mais vitimizaram os jovens do município, com a taxa de 160,0 por 100.000 habitantes entre 0 a 19 anos de idade. Os acidentes de trânsito perfizeram 97% deles, entre os quais 62% foram atropelamentos, 22% colisões e 16% outros tipos de acidentes de trânsito. Essas autoras mostram que, entre o conjunto das causas externas específicas, os acidentes de trânsito foram os que mais vitimizaram as faixas etárias de 5 a 9 anos e de 10 a 14 anos em ambos os sexos, em todas as áreas programáticas (AP) do Município do Rio de Janeiro, apresentando as maiores taxas de morbidade. Essas taxas

variaram, no ano de 1990, na faixa etária de 5 a 9 anos, de 87,3 por 100.000 habitantes na AP 3.1, que compreende os bairros de Bonsucesso, Penha e Ilha do Governador entre outros, a 339,7 por 100.000 habitantes na AP 3.3, que abrange, também os bairros de Madureira, Irajá e Guadalupe. Na faixa de 10 a 14 anos estas taxas foram 75,0 por 100.000 habitantes na AP 3.1 e 347,0 por 100.000 habitantes na AP 3.3.

DUARTE et al. (2002) destacam que os estados brasileiros com menores taxas de pobreza, menores níveis de alfabetização e maiores taxas de crescimento populacional no período de 1991 a 2000 apresentam os maiores riscos de mortes por acidentes de trânsito.

LIMA e VASCONCELOS (1999) ao analisarem a mortalidade por acidentes de trânsito no Brasil, chamam a atenção para o aumento das taxas de mortalidade nos períodos de aumento da atividade econômica, especificamente, em 1986 (Plano Cruzado) e 1994 (Plano Real).

Os agravos por violência, especialmente os acidentes de trânsito, têm constituído uma demanda significativa do atendimento das emergências Hospitalares (PINHEIRO, 1994). Este fato também constitui uma realidade internacionalmente reconhecida (MILLER, 1993), seja pelos custos que representam tais acidentes, seja pela complexidade do atendimento que demandam.

DESLANDES e SILVA (2000) ao analisarem os atendimentos de emergência para as causas externas em dois hospitais públicos do Município do Rio de Janeiro verificaram que dos 5.151 casos pesquisados, 11,8% atendimentos (610 casos) foram devidos aos acidentes de trânsito envolvendo veículos automotores. A faixa etária de maior ocorrência foi de 20 a 39 anos e 69,3% das vítimas de acidentes de trânsito eram

do sexo masculino. Os autores destacam que de um total de 23 óbitos por causas externas ocorridos no hospital 1, e 22 óbitos por causas externas ocorridos no hospital 2, respectivamente, 10 e 11 óbitos foram devidos às lesões provocadas por acidentes de trânsito.

LIMA (1996) defende que o número de feridos por acidentes de trânsito é, pelo menos, 14 vezes maior que o número de mortos e que entre os feridos, cerca de 20% ficam com seqüelas irreversíveis.

SCALASSARA et al. (1998) ao analisarem 65 vítimas fatais de acidentes de trânsito ocorridos em Maringá, em 1992, verificaram uma predominância do sexo masculino e da faixa etária de 20 a 49 anos (27 casos). Em relação aos dias da semana, os finais de semana (sábado e domingo) foram responsáveis por 50,8% dos óbitos registrados.

Das 628.976 internações por lesões e envenenamentos ocorridas no Brasil, em 2000, 18,2% foram devidas aos acidentes de transportes. Na faixa etária que compreende dos 20 a 29 anos a proporção de internações por lesões e envenenamentos, no Brasil, devida aos acidentes de transportes chegou a 23,1% do total de 125.072 internações (MINAYO et al., 2003).

Estudo de RODRIGUES (1995), sobre mortalidade hospitalar, demonstra que de todas as lesões no crânio atendidas em 1990, em Ribeirão Preto, a maioria deveu-se a acidentes de trânsito. Vale lembrar que a maioria das vítimas fatais de acidentes de trânsito não chega aos hospitais, morrendo no local do acidente ou no trajeto para os hospitais.

PORTUGAL e SANTOS (1991) destacam a necessidade de discutir o problema de trânsito no Brasil de forma ampliada, articulando atores e responsabilidades: das montadoras de carros, das empreiteiras de obras, dos empresários de ônibus, do poder público, das autoridades de trânsito, das comunidades, dos motoristas e pedestres. Os problemas do motorista alcoolizado, do avanço de sinal, do excesso de velocidade no perímetro urbano, das condições precárias de sinalização e pavimentação de ruas e estradas são corriqueiros, apesar dos muitos avanços decorrentes da implementação do Novo Código de Trânsito. Além disso, e o que é pior, as lesões decorrentes do trânsito continuam sendo percebidas como “acidentais” - eventos imprevisíveis, sendo determinados pelo destino.

LASCALA et al. (2000) ao analisarem a relação entre as taxas de acidentes de trânsito com pedestres e as características ambientais e demográficas na Cidade de São Francisco, no Estado da Califórnia, nos Estados Unidos da América, verificaram que uma maior densidade populacional e uma predominância de homens dentro de cada unidade geográfica (setor censitário), assim como, um maior fluxo de tráfego e uma maior concentração de pessoas desempregadas estavam associadas com maiores taxas de acidentes envolvendo pedestres. As menores taxas de acidentes envolvendo pedestres foram observadas nas unidades geográficas que apresentaram uma maior concentração de pessoas com idades até 15 anos e com níveis de escolaridade mais elevados. Ao analisarem a relação entre disponibilidade ao álcool e o envolvimento em acidentes de trânsito com pedestres, os autores verificaram que a quantidade de bares por quilometro de rodovia e a proporção de pessoas desempregadas, estão altamente associados com a ocorrência de acidentes de trânsito envolvendo pedestres que declararam ter consumido bebida alcoólica.

Estudo de JEGEDE (1988) buscando identificar os fatores que poderiam influenciar na variação espacial e temporal da ocorrência e fatalidade dos acidentes de trânsito no Estado de Oyo, na Nigéria, verificou através de um modelo de regressão múltiplo para o número de acidentes de trânsito que os fatores com maior contribuição para o modelo foram o número de estabelecimentos industriais, o tamanho da população, o comprimento das estradas municipais, o número de veículos registrados e o comprimento das estradas federais. Essas cinco variáveis explicavam pouco mais de 94% do modelo. O estudo identificou seis áreas de maior concentração de acidentes onde o autor sugere uma maior atenção dos programas de segurança. O autor utilizou o Índice Sazonal Específico, que é obtido através da razão entre o número de acidentes ocorridos e a média móvel centrada em 12, para verificar as variações sazonais, obtendo que os meses de março, setembro e dezembro foram os que apresentaram os maiores índices sazonais.

KMET et al. (2003) ao estudarem os acidentes de trânsito fatais de motoristas e ocupantes de veículos automotores, no período de 01/01/95 até 31/12/97, procuraram determinar as taxas de mortalidade por acidentes de trânsito em pequenas áreas da província de Alberta, no Canadá, avaliando as variações nessas taxas e buscando examinar o papel das variáveis explicativas em um modelo para a variação nas taxas de acidentes de trânsito através das áreas geográficas. A província de Alberta foi dividida em cinco pequenas áreas, sendo duas delas urbanas e as três restantes, rurais. Foram considerados no estudo, somente os motoristas e os ocupantes de veículos automotores que morreram em decorrência do evento. Em Alberta, considera-se como morte por acidente de trânsito aquela que ocorre no local do evento ou em decorrência do mesmo em até 30 dias. Os autores encontraram que uma taxa de 12,2 óbitos por acidentes de

trânsito em 100.000 habitantes, com uma sobremortalidade masculina maior que duas vezes a feminina (17,0 e 7,4, respectivamente). A maior incidência, em ambos os sexos, ocorreu na faixa etária de 15 a 24 anos. Para o ajuste do modelo de regressão de Poisson multivariado para avaliar a importância das variáveis explicativas na variação das taxas de acidentes de trânsito através das áreas geográficas, foram utilizadas como variáveis explicativas, sexo, faixa etária, densidade populacional, taxa de desemprego, nível de educação (proporção de pessoas com menos de 9 anos de estudo), etnia (proporção de aborígenes). Contudo, o ajuste não foi possível de ser realizado por problemas de multicolinearidade entre as variáveis. Os autores concluem que os acidentes de trânsito fatais de motoristas e ocupantes de veículos automotores são distribuídos de forma desigual na província de Alberta. Chamam a atenção das taxas observadas nas áreas rurais serem cinco vezes superiores àquelas observadas nas áreas urbanas e ressaltam que tal discrepância pode estar associada às variações de exposição ao risco, as características do evento, aos hábitos dos motoristas, as condições ambientais e a disponibilidade de serviços de saúde capacitados nessas áreas.

Com o objetivo de formular modelos de predição do número de acidentes esperados em junções urbanas e ligações de estradas, GREIBE (2003) utilizou 1.024 junções e 142 km de ligações de estradas em áreas urbanas da Dinamarca. As junções foram classificadas em quatro tipos: a) com três braços e sinalizada; b) com três braços e não sinalizada; c) com quatro braços e sinalizada; d) com quatro braços e não sinalizada. Os 142 km de estradas foram separados em 314 segmentos com comprimento médio de 450 metros e desvio padrão de 285 metros. Os modelos buscaram identificar fatores que afetam a segurança nas estradas em relação aos pontos de maior ocorrência de acidentes através da incorporação de variáveis explicativas. O

autor utilizou técnicas de modelos lineares generalizados para relacionar a frequência dos acidentes e as variáveis explicativas. Apesar de ter ressaltado que o indicado seria utilizar a Distribuição Binomial Negativa, pois permite uma maior discrepância nos dados, o autor assumiu que a contagem do número de acidentes segue uma Distribuição de Poisson, devido a sua simplicidade, uma vez que nessa distribuição, média e variância são iguais. Embora a utilização dessa distribuição não afete as estimativas dos coeficientes, ela pode causar uma subestimação dos erros padrões. O processo de modelagem iniciou com a inclusão de todas as variáveis no modelo, sendo excluídas, uma a uma, as variáveis não significativas. A variação total no número de acidentes consiste de uma parte aleatória (descrita pela Distribuição de Poisson) e por uma parte sistemática. As variáveis mais poderosas nos modelos foram as que descreveram o fluxo de tráfego de automóveis, particularmente para as junções. As demais variáveis explicativas nos modelos de junção não melhoraram o percentual de explicação consideravelmente. Para as ligações de estrada, porém, as variáveis explicativas adicionais que descrevem o ambiente da estrada, o número de estradas laterais secundárias, as facilidades de estacionamento (acostamentos) e a velocidade máxima provaram ser variáveis significantes e importantes para predizer o número de acidentes.

MIAOU (1994) analisou o desempenho do modelo de regressão de Poisson e do modelo de regressão Binomial Negativa para estabelecer a relação entre acidentes de caminhão e características geométricas de seções de estrada. Foram considerados três tipos de modelos: regressão de Poisson, regressão de Poisson inflacionado de zeros, e regressão Binomial Negativa. O autor utilizou o método de máxima verossimilhança para estimar os parâmetros dos modelos. Os modelos e estimativas foram avaliados segundo o cálculo dos parâmetros de regressão, a adequação do ajuste global, a

freqüência relativa estimada do envolvimento de acidente de caminhão por seções de estrada, a sensibilidade para a inclusão de seções de estrada curtas, e número total de envolvimento de acidentes de caminhão. O autor sugere que inicialmente o modelo de regressão de Poisson deva ser usado para se estabelecer a relação entre as características das rodovias e os acidentes. Caso seja encontrada uma superdispersão nos dados, o autor sugere que o modelo de regressão Binomial Negativa e o modelo de regressão de Poisson inflacionado de zeros poderiam ser explorados, sendo este um forte candidato quando a distribuição apresenta um grande número de observações com valores zero, embora sua interpretação possa ser mais difícil.

CHANG (2005) comparou a performance entre um modelo de regressão Binomial Negativa e um modelo de Rede Neural Artificial. Este autor demonstrou que a utilização de modelos de Redes Neurais Artificiais é um método alternativo consistente para analisar a freqüência de acidentes em rodovias, uma vez que a sua utilização não pressupõe qualquer relação entre as variáveis dependente e independente que seria necessária estabelecer em um modelo de regressão.

Com o objetivo principal de analisar as propriedades estatísticas de dois modelos de regressão linear convencionais e dois modelos de regressão de Poisson, MIAOU e LUM (1993) investigaram a relação entre os acidentes de trânsito e as características geométricas de rodovias. O estudo demonstrou que nos modelos de regressão linear convencionais falta a propriedade para descrever adequadamente a aleatoriedade tipicamente relacionada aos eventos de acidentes de trânsito. Assim, esses modelos não são apropriados para descrever os acidentes de trânsito, e suas estatísticas de teste são questionáveis. Os modelos de regressão de Poisson apresentaram a maioria das propriedades estatísticas desejáveis para descrever essas relações. Contudo, os autores

destacam, que se os dados apresentarem superdispersão significativa, o uso dos modelos de regressão de Poisson pode superestimar ou subestimar a probabilidade de acidentes de trânsito em estradas e assim distribuições de probabilidades mais gerais, tais como a Binomial Negativa, devem ser utilizadas.

FRIDSTROM et al. (1995) mediram a contribuição da aleatoriedade, exposição, clima e luz do dia, na variação do número de acidentes em estradas. Eles declararam que a formulação do modelo de regressão de Poisson generalizado para o número de acidentes permite a decomposição da variação total da variável dependente em uma parte devido à variação aleatória normal, que não é explicável, e em outra parte devido a fatores sistemáticos, causais.

SHANKAR et al. (1995) analisaram a ocorrência de acidentes na Rodovia Interestadual 90 (I-90) utilizando as características da própria via, do clima e também de outros efeitos sazonais, com o objetivo de focar os determinantes não-comportamentais relacionados ao risco de acidentes. Para tal utilizaram o modelo de regressão de Poisson e modelo de regressão Binomial Negativa. O modelo de Poisson foi utilizado quando a frequência de acidentes não apresentou superdispersão significativa, o que violaria um dos pressupostos para a sua utilização. O modelo Binomial Negativa foi utilizado quando a superdispersão foi verificada. Além da frequência de acidente global em seções da rodovia, os autores modelaram a frequência de tipos específicos de acidentes e concluíram que os modelos de regressão separados para cada tipo específico de acidentes têm um maior poder explicativo.

O modelo de regressão Binomial Negativa foi utilizado para modelar a frequência de acidentes de trânsito por ABDEL-ATY e RADWAN (1999). Os autores

analisaram os dados de acidentes de trânsito de um período de 3 anos, compreendidos de 1992 a 1994, totalizando 1606 acidentes ocorridos na Rodovia Estadual 50 (SR 50), que liga o Leste ao Oeste da Flórida Central, passando pelo centro de Orlando. Essa rodovia apresenta áreas rurais e urbanas, e o número total de pistas varia entre duas, quatro e seis.

O modelo destacou a importância das variáveis fluxo de tráfego diário, grau de curvatura (curvas), largura das pistas, largura dos acostamentos e largura dos canteiros centrais, área urbana e rural, comprimento das seções de estrada, na frequência de ocorrência dos acidentes. Também foram desenvolvidos modelos utilizando a distribuição Binominal Negativa para descrever o número de acidentes segundo as características demográficas do motorista (idade e gênero). Os resultados mostraram que o volume de tráfego mais intenso, a velocidade maior, a largura estreita da pista, o maior número de pistas, a existência de seções urbanas da estrada, a largura estreita dos acostamentos e a largura reduzida do canteiro central aumenta a probabilidade para a ocorrência de acidentes.

Em relação aos motoristas, os autores verificaram que as mulheres se envolvem em mais acidentes que os homens quando o volume de tráfego é mais intenso, a largura do canteiro central é reduzida, a largura das pistas é mais estreita, e o número de pistas é maior. Os homens, além de apresentar as mesmas variáveis encontradas para as mulheres para descrever o número de acidentes de trânsito, têm, por sua vez, maior tendência a se envolverem em acidentes de trânsito enquanto aceleram. Os modelos também indicaram que os motoristas mais jovens e os mais velhos sofrem mais acidentes que os motoristas de meia-idade quando o volume de tráfego é mais intenso e os acostamentos e canteiros centrais têm larguras mais reduzidas. Os motoristas mais

jovens têm maior tendência a se envolverem em acidentes em curvas mais acentuadas (curvas mais fechadas) em estradas e enquanto aceleram do que os motoristas mais idosos. Os autores concluem que o modelo de regressão Binomial Negativa é superior ao modelo de Poisson, e acrescentam à literatura a dimensão de incluir os efeitos da idade e do gênero dos condutores na modelagem de acidentes de trânsito.

AMOROS et al. (2003) compararam a segurança no trânsito de diferentes regiões da França com o objetivo de investigar se as diferenças observadas poderiam ser explicadas pelas diferenças da distribuição dos tipos de rodovias e pelas diferenças nas características sócio-econômicas das regiões. Para isso analisaram a contagem do número de acidentes com um modelo de regressão binomial negativa, uma vez que essas contagens apresentaram superdispersão, violando uma condição para a utilização do modelo de Poisson. Os autores verificaram que o modelo apresentou como variáveis significativas, a proporção de pessoas com idades entre 35 e 64 anos, a razão entre o número de novas licenças de motorista e o número de veículos, e a razão entre o número de instrutores de direção e o número de veículos.

Observa-se então que grande parte dos trabalhos, acerca dos acidentes de trânsito descritos na literatura, apontam para a utilização de modelos que visam identificar os fatores que podem influenciar na variação espacial e temporal da ocorrência e fatalidade desse evento.

3. BASES TEÓRICAS

3.1. Análise Espacial

A análise espacial é um conjunto de métodos quantitativos utilizados com o objetivo de estudar fenômenos distribuídos no espaço (BAILEY e GATRELL, 1995). Os dados espaciais são as medidas e contagens, efetuadas do fenômeno em estudo, que se distribuem em locais específicos. A análise espacial tem como um dos principais fundamentos que os dados obtidos em locais próximos (regiões próximas entre si) são mais correlacionados que os dados obtidos em regiões distantes entre si.

Os métodos utilizados na análise espacial dividem-se basicamente em dois grupos: análise exploratória e modelagem. Na análise exploratória utilizam-se procedimentos para a descrição, reconhecimento e exploração do processo espacial a partir das informações obtidas. Na modelagem busca-se a obtenção de explicações para os padrões observados e suas possíveis relações.

O primeiro passo na análise exploratória de dados espaciais é a construção de mapas coropléticos (CRESSIE, 1991). Esses mapas são utilizados para a visualização dos dados. A partir da visualização dos padrões, realizada pelos mapas coropléticos, são analisadas as propriedades de primeira ordem (tendência) e de segunda ordem (dependência espacial). Para a análise das propriedades de primeira ordem é utilizada a Média Móvel Espacial (RIPLEY, 1981).

A Média Móvel Espacial é obtida através da relação entre o número de ocorrências do evento em cada área (y_i) e a matriz de proximidade de áreas (w_{ij}), onde i e j são as áreas:

$$\hat{\mu}_i = \frac{\sum_{j=1}^n w_{ij} y_j}{\sum_{j=1}^n w_{ij}} \quad (1)$$

Quando a matriz de proximidade de áreas é normalizada, $\hat{\mu}_i$ reduz-se a:

$$\hat{\mu}_i = \sum_{j=1}^n w_{ij}^* y_j \quad (2)$$

Então, a média móvel espacial da área i ($\hat{\mu}_i$) é dada pela média dos vizinhos da área i , aqui representados por j .

Para analisar as propriedades de segunda ordem (dependência espacial) são utilizados o I de Moran e o C de Geary. O I de Moran e o C de Geary são técnicas utilizadas para analisar as propriedades de segunda ordem para dados de áreas, uma vez que, utiliza, de forma mais adequada, a matriz de proximidades W . São medidas de correlação espacial que avaliam o quanto o valor medido, em uma determinada área depende de valores de áreas vizinhas. São dadas por:

a) I de Moran

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right) \sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

b) C de Geary

$$C = \frac{(n-1) \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - y_j)^2}{2 \left(\sum_{i=1}^n \sum_{j=1}^n w_{ij} \right) \sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

A interpretação do I de Moran é semelhante à interpretação de um coeficiente de correlação não-espacial, ou seja, quanto mais próximo de 0 indica falta de associação espacial e para valores próximos de 1 indica forte associação. Por outro lado, a interpretação do C de Geary é o oposto, ou seja, quanto mais próximo de 1 há indícios de que não existe associação espacial e quanto mais próximo de 0 mais forte é a associação entre localidades vizinhas.

Para a avaliação desses coeficientes de correlação é utilizado o teste de independência espacial que tem como hipótese nula (H_0) a completa aleatoriedade espacial. Assim, são estimadas a média e a variância da distribuição aproximada e a partir daí é utilizada a variável normal padronizada Z para a realização dos testes.

Ou seja, para o I de Moran

$$\frac{\hat{I} - E(I)}{\sqrt{\text{Var}(I)}} = Z \sim N(0, 1) \quad (5)$$

onde:

$$E(I) = -\frac{1}{(n-1)} \quad (6)$$

$$\text{Var}(I) = \frac{n^2(n-1)S_1 - n(n-1)S_2 - 2S_0^2}{(n+1)(n-1)^2 S_0^2} \quad (7)$$

Para C de Geary

$$\frac{\hat{C} - E(C)}{\sqrt{\text{Var}(C)}} = Z \sim N(0, 1) \quad (8)$$

onde:

$$E(C) = 1 \quad (9)$$

$$\text{Var}(C) = \frac{(2S_1 + S_2)(n-1) - 4S_0^2}{2(n+1)S_0^2} \quad (10)$$

Para os dois índices:

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij} \quad (11)$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1, j \neq i}^n (w_{ij} + w_{ji})^2 \quad (12)$$

$$S_2 = \sum_k \left(\sum_j w_{kj} + \sum_i w_{ik} \right)^2 \quad (13)$$

Uma maneira adicional de se visualizar a dependência espacial é dada pelo diagrama de espalhamento de Moran, pois permite analisar o comportamento da variabilidade espacial.

O diagrama de espalhamento de Moran é um gráfico bidimensional, dividido em quatro quadrantes, onde são plotados os valores normalizados do fenômeno numa área com a média dos seus vizinhos, para todas áreas (figura 1). No primeiro quadrante do diagrama de espalhamento de Moran são plotados os valores positivos e as médias positivas e no segundo quadrante são plotados os valores negativos e as médias negativas. Os pontos localizados no primeiro e segundo quadrantes possuem associação espacial positiva, de tal forma que uma determinada área possui vizinhos com comportamentos semelhantes aos dela. O terceiro quadrante possui valores positivos e médias negativas, e o quarto quadrante possui valores negativos e médias positivas. Os pontos localizados no terceiro e quarto quadrantes possuem associação espacial negativa, de tal forma que uma determinada área possui vizinhos com comportamentos diferentes dos dela.

Quando os valores observados nas áreas não estão padronizados, a determinação dos quadrantes é feita da seguinte maneira: o ponto que divide o eixo das abscissas em duas partes é obtido através da média das médias dos vizinhos e o ponto que divide o eixo das ordenadas em duas partes é obtido através da média dos valores observados nas áreas abrangidas pelo estudo.

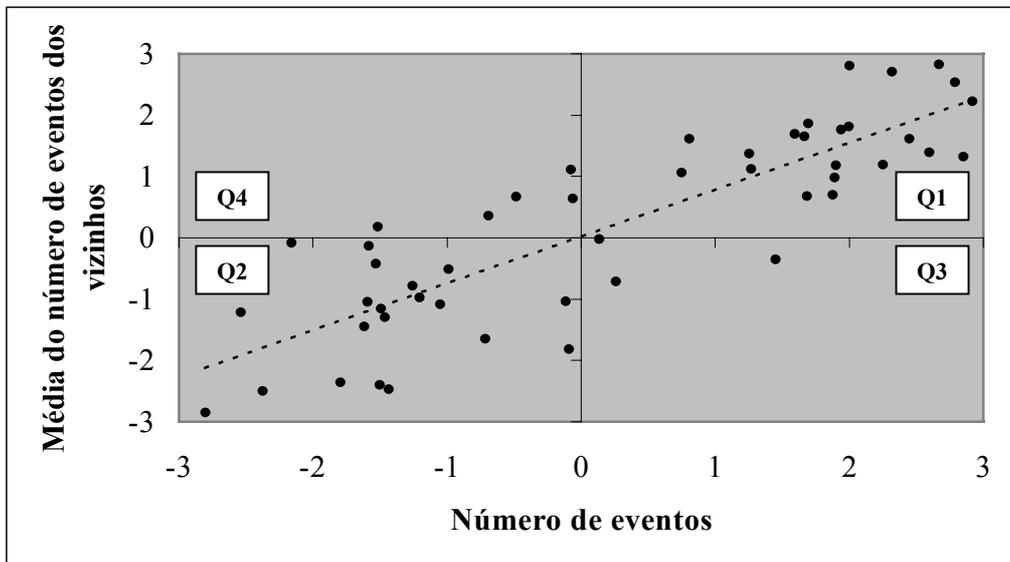


Figura 1: Diagrama de espalhamento de Moran.

3.2. Modelos Lineares Generalizados

O processo de modelagem consiste na busca de um modelo, o mais simples possível, que descreva bem os dados observados. Durante esse processo considera-se, na maioria dos casos, que o comportamento da variável resposta pode ser representado por dois componentes: um componente sistemático e um aleatório (KRZANOWSKI, 1998; CORDEIRO, 1986; KOTZ e JOHNSON, 1983; DOBSON, 1983). O componente sistemático é estabelecido durante o planejamento do experimento, resultando em modelos de regressão, de análise de variância e de análise de covariância. O componente aleatório é estabelecido logo que são definidas as medidas a serem realizadas.

No modelo de regressão linear clássico (HOFFMANN e VIEIRA, 1998; NETER *et al.*, 1996; LARSON, 1982; DRAPER e SMITH, 1981), defini-se:

$$Y_i = \mu_i + \varepsilon_i \quad (14)$$

onde:

Y_i é a variável resposta;

$\mu_i = E(Y) = X\beta$ representa o componente sistemático também denominado previsor linear;

ε_i representa o componente aleatório.

Para um conjunto de n observações onde foram medidos os valores de Y e de K variáveis (X_1, X_2, \dots, X_k) a serem usadas como variáveis explicativas do modelo, temos:

$$\tilde{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \text{ que representa o vetor de observações;}$$

$$\tilde{X} = \begin{bmatrix} 1 & X_{11} & X_{21} & \cdots & X_{k1} \\ 1 & X_{12} & X_{22} & \cdots & X_{k2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_{1n} & X_{2n} & \cdots & X_{kn} \end{bmatrix}, \text{ que representa a matriz de observações para as}$$

variáveis explicativas;

$$\tilde{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \text{ que representa o vetor de coeficientes;}$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}, \text{ que representa o vetor de erros.}$$

Tem-se, então, as seguintes pressuposições para o modelo:

- a) a variável resposta, também chamada de variável dependente, (Y_i) é função linear das variáveis independentes $(X_{ki}, \text{ para } k = 1, \dots, K)$;
- b) as variáveis independentes não são consideradas variáveis aleatórias e seus valores são fixos e conhecidos;
- c) $E(\varepsilon_i) = 0$, para $i = 1, \dots, n$;

- d) os erros são homocedásticos, ou seja, $E(\varepsilon_i^2) = \sigma^2$;
- e) os erros não são correlacionados entre si, ou seja, $E(\varepsilon_i \varepsilon_j) = 0 \quad \forall i \neq j$;
- f) o componente aleatório tem distribuição Normal com média zero e variância σ^2 ;
 $\varepsilon_i \sim N(0, \sigma^2), i = 1, \dots, n$.

Assim, verifica-se que, nos modelos de regressão linear clássicos, os erros têm distribuição Normal e a variável resposta (dependente) é expressa como uma combinação linear de efeitos fixos (da parte sistemática) e dos efeitos aleatórios. Deve-se ressaltar que a estrutura aditiva entre o componente sistemático e o componente aleatório, assim como a restrição da componente aleatória ter distribuição Normal e ter comportamento homocedástico, ou seja variâncias iguais ($V(Y_i) = \sigma^2, \forall i$), nem sempre são satisfeitas. Além disso, não é adequado supor que a relação entre a média da variável resposta e o preditor linear deve estar unicamente restrita à estrutura $\mu = E(Y) = X\beta$.

Os modelos lineares generalizados apresentam, como uma das principais características, a capacidade de unificar diversas técnicas estatísticas que, geralmente, são estudadas separadamente. Este tipo de modelagem é utilizado, principalmente, quando um ou mais pressupostos dos modelos de regressão linear clássicos não são satisfeitos. Além disso, admiti-se que as variáveis respostas possuam distribuição diferente da Normal e que sua média se expresse como diferentes funções de covariáveis e de seus respectivos coeficientes.

Desta forma, os modelos lineares generalizados são utilizados quando, para n observações, se tem uma única variável aleatória Y_i considerada, como variável resposta, cujo comportamento esteja associado a um conjunto de variáveis explicativas X_1, \dots, X_K .

O componente aleatório, o componente sistemático e a função de ligação são os três componentes que caracterizam os modelos lineares generalizados.

No componente aleatório do modelo, as variáveis respostas (Y_1, \dots, Y_n) são aleatórias e independentes e têm uma distribuição pertencente à família exponencial na forma canônica com média μ_i e parâmetro de dispersão ou escala dado por $a_i(\phi) = \frac{\phi}{w_i}$,

sendo w_i o peso a priori e $\phi > 0$, ou seja,

$$f(y_i; \theta, \phi) = \exp \left\{ \frac{1}{a_i(\phi)} [y_i \theta_i - b(\theta_i)] + c(y_i; \phi) \right\} \quad (15)$$

em que $b(\cdot)$ e $c(\cdot)$ são funções conhecidas, θ_i é chamado de parâmetro canônico. Além disso, $E(Y_i) = \mu_i = b'(\theta_i)$ e $\text{Var}(Y_i) = a_i(\phi) b''(\theta_i)$, onde b' e b'' representam, respectivamente, a primeira e a segunda derivadas.

O quadro 1 apresenta alguns tipos de distribuições pertencentes à família exponencial na forma canônica segundo o tipo de variável resposta e o quadro 2, os identificadores da família exponencial dessas distribuições.

Quadro 1: Variáveis respostas e distribuição pertencente à família exponencial na forma canônica do componente aleatório dos modelos lineares generalizados.

Tipo de variável resposta	Distribuição pertencente à família exponencial na forma canônica
Contínua	Distribuição Normal
	Distribuição Gama
	Distribuição Normal Inversa
Proporção	Distribuição Binomial
Contagem	Distribuição de Poisson
	Distribuição Binomial Negativa

Quadro 2: Identificadores da família exponencial das distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.

Distribuição	$a(\phi)$	θ	$b(\theta)$	$c(y; \phi)$
Normal $N(\mu, \sigma^2)$	σ^2	μ	$\frac{\theta^2}{2}$	$-\frac{1}{2} \left[\frac{y^2}{\sigma^2} + \ln(2\pi\sigma^2) \right]$
Gama $G(\mu, \nu)$	ν^{-1}	$-\frac{1}{\mu}$	$-\ln(-\theta)$	$\nu \ln(\nu y) - \ln y - \ln \Gamma(\nu)$
Normal Inversa $IG(\mu, \sigma^2)$	σ^2	$-\frac{1}{2\mu^2}$	$-(-2\theta)^{1/2}$	$-\frac{1}{2} \left[\ln(2\pi\sigma^2 y^3) + \frac{1}{\sigma^2 y} \right]$
Binomial $B(m, \pi)$	1	$\ln\left(\frac{\pi}{1-\pi}\right)$	$m \ln(1 + e^\theta)$	$\ln\left(\frac{m}{y}\right)$
Poisson $P(\mu)$	1	$\ln \mu$	e^θ	$-\ln y!$
Binomial Negativa $BinNeg(\mu, k)$	1	$\ln\left(\frac{\mu}{\mu+k}\right)$	$-k \ln(1 - e^\theta)$	$\ln\left[\frac{\Gamma(k+y)}{\Gamma(k)y!}\right]$

No componente sistemático dos modelos lineares generalizados as variáveis explicativas entram na forma de uma combinação linear de seus efeitos dando origem ao preditor linear η_i , que é um vetor de dimensão $n \times 1$. Assim,

$$\eta_i = X_i' \beta \quad (16)$$

onde X_i é a matriz de covariáveis e β o vetor de parâmetros desconhecidos a serem estimados.

O terceiro componente dos modelos lineares generalizados é a função de ligação. A função de ligação $g(\cdot)$ é uma função conhecida e diferenciável, que tem o objetivo de relacionar o componente aleatório (μ_i) e o componente sistemático (η_i), ou seja, relaciona a média ao preditor linear, ou seja,

$$g(\mu_i) = \eta_i = X_i' \beta \quad (17)$$

Quando a função de ligação escolhida é tal que $g(\mu_i) = \theta_i$, ou seja, o preditor linear modela diretamente o parâmetro canônico, a função de ligação é chamada de função de ligação canônica. Assim, a função de ligação pode ser obtida, diretamente, da distribuição de probabilidade na forma da família exponencial. O modelo de regressão logístico, o modelo log-linear assim como o modelo de regressão Normal são casos particulares dos modelos lineares generalizados.

As funções de ligação canônicas para as distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa são apresentadas no quadro 3.

Quadro 3: Funções de ligação canônicas das distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.

Distribuição	Tipo de função	Ligação canônica
Normal	Identidade	$\eta = \mu$
Gama	Recíproca	$\eta = \frac{1}{\mu}$
Normal Inversa	Recíproca ²	$\eta = \frac{1}{\mu^2}$
Binomial	Logística	$\eta = \ln\left(\frac{\pi}{1-\pi}\right) = \ln\left(\frac{\mu}{m-\mu}\right)$
Poisson	Logarítmica	$\eta = \ln \mu$
Binomial Negativa	Logarítmica	$\eta = \ln\left(\frac{\mu}{\mu + k}\right)$

A escolha da distribuição de probabilidade da variável resposta, da matriz de variáveis explicativas e da função de ligação é de extrema importância para a construção do modelo linear generalizado.

Para o ajuste de um modelo linear generalizado torna-se necessário que o pesquisador estabeleça uma estrutura de análise de tal forma que possibilite a identificação do modelo, a estimação dos parâmetros do modelo, a avaliação do ajuste do modelo e a avaliação da adequação do modelo, visando assim que o mesmo possa representar, o mais próximo possível, a realidade que busca retratar.

A identificação do modelo depende, basicamente, de três fatores: a escolha das variáveis (dependente e independentes) que serão, inicialmente, consideradas no modelo; do tipo de relação que existe entre essas variáveis; e também da escala de medidas dessas variáveis.

Na escolha das variáveis que serão, inicialmente, utilizadas no modelo, devem ser consideradas somente aquelas que, de acordo com a avaliação dos especialistas, possuem alguma ligação lógica com o problema em questão. Sendo assim, as variáveis devem ser escolhidas de acordo com a sua relevância para estudo. Além disso, deve-se tomar cuidado na inclusão de variáveis que apresentem autocorrelação serial ou dependência entre si, para evitar problemas na hora de escolha do melhor modelo.

A escala de medida das variáveis é um outro fator importante na definição do modelo de regressão. Uma boa escala deve combinar homocedasticidade, normalidade dos erros e efeitos sistemáticos aditivos.

Para o ajuste de um modelo linear generalizado é necessário a determinação do vetor ($\hat{\beta}$) de estimativas dos parâmetros. Esses coeficientes são estimados a partir dos dados observados. A avaliação do modelo é realizada utilizando-se uma medida de qualidade do ajuste que considera a relação entre os dados observados e os valores estimados pelo modelo. Na determinação do vetor ($\hat{\beta}$) de estimativas dos parâmetros dos modelos lineares generalizados são utilizados os Estimadores de Máxima Verossimilhança. Essas estimativas são obtidas através do método iterativo de Newton-Raphson, ou do método de escore de Fisher ou do método de mínimos quadrados ponderados.

Após o ajuste do modelo é necessário avaliar a sua qualidade. O objetivo é determinar quantos e quais termos são necessários para uma boa descrição dos dados. Um número muito grande de variáveis explicativas, embora possa descrever muito bem os dados, implica, também, em complexidade da interpretação. Por outro lado, um número muito pequeno de variáveis explicativas pode não ser suficiente para descrever bem os dados, embora possua fácil interpretação. Sendo assim, o que na realidade se deseja é um modelo o mais parcimonioso possível.

O modelo mais completo, também chamado de modelo saturado, é aquele que possui o mesmo número de observações e parâmetros. Esse tipo de modelo atribui toda a variação ao componente sistemático, possuindo um ajuste perfeito e reproduzindo, assim, os próprios dados. Por esse motivo, o modelo saturado é não informativo, pois não resume os dados.

Na outra extremidade da escala de complexidade está o modelo nulo que possui um único parâmetro, representado por um mesmo valor para todas as observações (este valor é em geral a média das observações). Esse tipo de modelo atribui toda a variação ao componente aleatório.

Assim, o objetivo do processo de modelagem é estimar um modelo que apresente um pequeno conjunto de parâmetros, de tal forma que a distância entre os valores observados e os valores previstos esteja dentro de limites aceitáveis de erro.

Considerando $\hat{\theta}$ e $\tilde{\theta}$ como estimativas do parâmetro canônico para os modelos estimado e saturado, respectivamente, tem-se:

$$D = 2 \sum_{i=1}^n [y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i)] \quad (18)$$

A função D é denominada como Deviance ou função desvio do modelo e a razão $\frac{D}{\phi}$ é denominada escala da função desvio, onde ϕ é o parâmetro de dispersão. A função desvio (D) é uma medida de discrepância que utiliza o modelo saturado como base de medida do ajuste do modelo estimado. Logo, é uma medida da distância dos valores ajustados $\hat{\mu}$'s em relação aos dados observados y 's, ou seja, do modelo estimado em relação ao modelo saturado.

A avaliação da qualidade do modelo é obtida a partir dos resultados obtidos para a função desvio: grandes valores indicam uma inadequação do ajuste enquanto que, valores menores nesta escala indicam melhoria na qualidade do ajuste. A questão principal fica em decidir quais valores são pequenos e quais valores são grandes. A distribuição da função desvio é desconhecida, entretanto, pode ser aproximada para uma distribuição Qui-quadrado com $n-p$ graus de liberdade, onde n é número de observações e p é o número de parâmetros do modelo de interesse (modelo estimado). O teste consiste em comparar o valor obtido na escala da função desvio com a coordenada da função Qui-quadrado com $n-p$ graus de liberdade, considerando um nível de significância α pré-estabelecido. Assim, se $\chi_0^2 > \chi_{n-p}^2$ indica um ajuste ruim.

O quadro 4 apresenta a escala da função desvio para algumas distribuições de probabilidade da família exponencial.

Quadro 4: Função desvio para as distribuições Normal, Gama, Normal Inversa, Binomial, Poisson e Binomial Negativa.

Distribuição	Função Desvio
Normal	$\frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \hat{\mu}_i)^2$
Gama	$2v \sum_{i=1}^n \left[-\ln\left(\frac{y_i}{\hat{\mu}_i}\right) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} \right]$
Normal Inversa	$\frac{1}{\sigma^2} \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{y_i \hat{\mu}_i^2}$
Binomial	$2 \sum_{i=1}^n \left[y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) + (m_i - y_i) \ln\left(\frac{m_i - y_i}{m_i - \hat{\mu}_i}\right) \right]$
Poisson	$2 \sum_{i=1}^n \left[y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - \hat{\mu}_i) \right]$
Binomial Negativa	$2 \sum_{i=1}^n \left[y_i \ln\left(\frac{y_i}{\hat{\mu}_i}\right) - (y_i - k) \ln\left(\frac{y_i + k}{\hat{\mu}_i + k}\right) \right]$

Um importante ponto a ser destacado na modelagem estatística refere-se à decisão de quais variáveis explicativas podem ser retiradas ou não do modelo sem que isto prejudique significativamente o ajuste ao conjunto de dados. No modelo de regressão linear múltipla a verificação da hipótese, de existência de relação linear entre as variáveis explicativas (X) e a variável resposta (Y), é feita por meio de uma análise da variância.

Considera-se como hipótese nula (H_0) a não existência de relação linear entre as variáveis explicativas e a variável resposta, ou seja,

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_K = 0,$$

contra a hipótese alternativa de que pelo menos um coeficiente de parâmetro (β_i) seja diferente de zero ($\beta_i \neq 0$). A análise é realizada através da comparação do modelo que contempla toda a variação existente nos dados (modelo completo), com um modelo reduzido.

Nos Modelos Lineares Generalizados, o teste de ajuste do modelo também se dá através da comparação entre um modelo completo e um modelo reduzido, porém, devido à possibilidade de uso de qualquer distribuição da família exponencial, permitindo, assim, a utilização de variáveis que não possuam distribuição normal ou que não são ortogonais, a análise de variância torna-se inadequada para seleção do melhor modelo. Nestes casos opta-se pela análise da função desvio (deviance) para a seleção do modelo, que é uma generalização da análise de variância, que visa obter, a partir de uma seqüência de modelos encaixados, ou seja, a partir de uma seqüência de modelos onde cada modelo inclui mais termos que o modelo anterior, os efeitos de fatores, de covariáveis e de suas possíveis interações.

Assim, para se testar a adequação do modelo, nos Modelos Lineares Generalizados, utiliza-se a análise da função desvio, que consiste em verificar se $W = -2 \ln \Lambda_0$ é assintoticamente distribuída conforme uma Qui-quadrado com $(p - q)$ graus de liberdade ($W \sim \chi_{p-q}^2$).

Onde:

- Λ_0 é a razão entre \hat{L}_F e \hat{L}_R que são, respectivamente, a máxima verossimilhança para distribuição de Y no modelo com p variáveis

explicativas incluídas e no modelo reduzido com apenas q variáveis explicativas;

- p e q representam o número de parâmetros do modelo completo e do modelo reduzido, respectivamente.

Assim, se $W \leq \chi_{p-q;\alpha}^2$, pode-se considerar que há evidências, a um nível α de significância, que o modelo proposto está bem ajustado aos dados.

Considerando \hat{L}_S como a máxima verossimilhança do modelo saturado e dividindo \hat{L}_F e \hat{L}_R por \hat{L}_S , tem-se:

$$W = -2 \left(\ln \left(\frac{\hat{L}_R}{\hat{L}_S} \right) - \ln \left(\frac{\hat{L}_F}{\hat{L}_S} \right) \right) \quad (19)$$

Assumindo que $a_1(\phi) = \phi$ na definição do modelo e denotando a escala da função desvio dos modelos completo e reduzido por, respectivamente, $\frac{D_F}{\phi}$ e $\frac{D_R}{\phi}$ tem-se:

$$W = \frac{(D_R - D_F)}{\phi} \quad (20)$$

Quando ϕ não é conhecido, ele deve ser estimado por $\tilde{\phi} = \frac{D_F}{(n-p)}$ e então:

$$W = \frac{(D_R - D_F)/(p-q)}{D_F/(n-p)} \quad (21)$$

Se a hipótese nula (H_0) é verdadeira, o numerador e o denominador têm distribuição aproximadamente Qui-quadrado com, respectivamente, $(p - q)$ e $(n - p)$ graus de liberdade e as duas distribuições são independentes. Então, W tem distribuição F com $(p - q)$ e $(n - p)$ graus de liberdade, uma vez que, uma razão de variáveis Qui-quadrado tem aproximadamente uma distribuição F.

$$W \sim F_{p-q; n-p}$$

Na comparação de modelos encaixados, $M_1 \supset M_2 \supset M_3 \supset \dots \supset M_K$, a tabela de análise dos desvios é calculada a partir das funções desvio de cada modelo considerado (D_1, D_2, \dots), onde a função desvio D_1 tem mais parâmetros que a função desvio D_2 , a função desvio D_2 tem mais parâmetros que a função desvio D_3 e assim sucessivamente. Então, são tomadas as diferenças sucessivas entre os modelos ($D_2 - D_1, D_3 - D_2, \dots$) e o teste consiste em verificar se essas diferenças entre os modelos são significativas no ajuste ao conjunto de dados.

O processo de modelagem é uma aproximação da realidade que consiste, basicamente, em representá-la de uma forma simplificada. Esse processo envolve suposições a respeito da variável resposta e da sua relação com as variáveis explicativas. Desta forma, a adequação do modelo, ou seja, a capacidade de identificar quais serão os valores da variável resposta dada a ocorrência das variáveis explicativas, pode ser verificada através da diferença entre os valores esperado e observado, e esta diferença é denominada resíduo.

Os principais tipos de resíduos utilizados em modelos lineares generalizados encontram-se no quadro 5.

Quadro 5: Principais tipos de resíduos utilizados em Modelos Lineares Generalizados.

Tipo de Resíduo	Fórmula
Resíduo Ordinário	$r_i = y_i - \hat{\mu}_i$
Resíduo de <i>Pearson</i> Padronizado	$r_{i(P)} = \frac{y_i - \hat{\mu}_i}{\sqrt{[(1 - h_i)V(y_i)]}}$
Resíduo da Função Desvio Padronizado	$r_{i(D)} = \frac{\text{senal}(y_i - \hat{\mu}_i)\sqrt{D_i}}{\sqrt{(1 - h_i)}}$
Resíduo de Máxima Verossimilhança	$r_{i(L)} = \text{senal}(y_i - \hat{\mu}_i)\sqrt{h_i r_{i(P)}^2 + (1 - h_i)r_{i(D)}^2}$
Distância de Cook	$C_i = r_{i(L)} \sqrt{\frac{(n - p)h_i}{p(1 - h_i)}}$

Onde h_i é a diagonal principal da matriz $H = W^{1/2}X(X'WX)^{-1}X'W^{1/2}$, sendo que W é a matriz de pesos obtida através do estimador de máxima verossimilhança.

Assim, são utilizados os métodos de diagnóstico do modelo, que visam verificar a existência de grandes afastamentos das suposições feitas em relação ao modelo, como por exemplo, a falta de normalidade dos erros e se existem evidências de heterocedasticidade (variância não constante).

Para analisar a suposição de normalidade dos erros foi utilizado o Gráfico Normal de Probabilidades (*Normal Probability Plot*) que para sua construção, primeiramente, ordenam-se os resíduos padronizados em ordem crescente e plota-se o k-ésimo resíduo ordenado *versus* a sua probabilidade acumulada dos resíduos (P_k) dada por:

$$P_k = \frac{k - \frac{1}{2}}{N} \quad (22)$$

onde:

k é o número de ordem do resíduo;

N é o número total de resíduos.

Quando a variabilidade aumenta à medida que aumenta a magnitude das observações, é sinal que a variância não é constante. Sendo assim, para se verificar a suposição de homocedasticidade dos resíduos foi utilizado gráfico dos resíduos padronizados *versus* os valores estimados. Esse tipo de gráfico, quando não apresenta nenhum tipo de tendência, é um indicativo que a relação funcional variância/média proposta para as observações é satisfatória.

3.2.1. Modelos de Regressão de Poisson

Os modelos de regressão de Poisson são utilizados quando a variável resposta retrata uma contagem. Este tipo de modelo linear generalizado, que possui como função de ligação a função logarítmica, é muito empregado quando a variável aleatória é o número de ocorrências do evento em um intervalo (que pode ser definido nas escalas de tempo, distância, área, volume ou outra unidade análoga).

O modelo de Poisson tem os seguintes pressupostos:

- a) as observações são independentes;
- b) o logaritmo das contagens varia linearmente com as covariáveis;
- c) o número de ocorrências tem média e variância iguais em cada nível das covariáveis.

Assim,

$$Y_i \sim \text{Poisson}(\mu_i), \text{ para } i = 1, \dots, n$$

e sua função de probabilidade é dada por:

$$f(y_i; \mu_i) = \frac{\mu_i^{y_i} e^{-\mu_i}}{y_i!}, \quad \mu_i > 0, \quad y_i = 0, 1, 2, \dots \quad (23)$$

No caso da distribuição de Poisson cujo modelo possui função de ligação logarítmica costuma-se utilizar a correção $y_i = y_i + \frac{1}{2}$, pois se $y_i = 0$, o processo não

pode ser iniciado, uma vez que o logaritmo natural não pode ser calculado para valores menores ou iguais a zero. Assim, substitui-se y por $y + c$ de tal forma que $E[g(Y + c)]$ seja o mais próximo possível de $g(\mu)$.

Destaca-se que o uso da função de ligação logarítmica faz com que os valores ajustados de μ_i permaneçam no intervalo $[0, \infty[$. Esse tipo de modelo é também conhecido como modelo log-linear e a sua expressão é:

$$\ln(\mu_i) = \eta_i = X_i' \beta = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} \quad (24)$$

Assim, a expressão para a variável resposta é dada por:

$$y_i = e^{\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik}} \quad (25)$$

Deve-se tomar cuidado na análise de dados, principalmente, quando a variável de interesse é uma variável aleatória discreta, como no caso de uma contagem, com o problema da superdispersão (HU *et al.*, 1998, BERHANU, 2004, CASTILLO e PÉREZ CASANY, 2005). No caso do modelo de regressão de Poisson, a superdispersão violaria o seguinte pressuposto:

$$E(Y_i) = \text{Var}(Y_i) = \mu_i$$

De forma que:

$$\text{Var}(Y_i) > \mu_i$$

Ou seja, a média e a variância não seriam iguais e assim a distribuição de Poisson não poderia ser utilizada.

Não considerar a superdispersão na análise dos dados pode acarretar a estimação incorreta dos erros padrões (super ou subestimando-os), ocasionando, assim, uma avaliação errônea da significância dos parâmetros de regressão.

3.2.2. Modelos de Regressão Binomial Negativa

O modelo de regressão Binomial Negativa é indicado para dados de contagem quando se observa superdispersão. O fenômeno da superdispersão ocorre quando a variável resposta tem distribuição de Poisson, contudo a variância é maior do que a sua média. Neste caso, assume-se que $Y_i \sim \text{Poisson}(\theta_i)$ e que os θ_i 's são variáveis aleatórias com média μ_i e variância σ_i^2 , ou seja, $E(\theta_i) = \mu_i$ e $\text{Var}(\theta_i) = \sigma_i^2$. Considerando $\theta_i \sim \Gamma(k, \lambda_i)$, tem-se uma distribuição Binomial Negativa para os Y_i 's.

Assim,

$$Y_i \sim \text{BinNeg}(\mu_i, k), i = 1, \dots, n, k > 0$$

com média e variância dadas por:

$$E(Y_i) = \frac{k}{\lambda_i} = \mu_i \quad \text{e} \quad \text{Var}(Y_i) = \mu_i + \frac{\mu_i^2}{k}$$

função de probabilidade dada por:

$$f(y; \mu, k) = \frac{\Gamma(k+y)}{\Gamma(k)y!} \frac{\mu^y k^k}{(\mu+k)^{k+y}} \quad (26)$$

e risco relativo para cada um das variáveis explicativas do modelo dado pela exponencial do coeficiente, ou seja:

$$RR_j = e^{(\beta_j)} \quad (27)$$

onde β representa o coeficiente da variável j significativa no modelo.

4. MATERIAL E MÉTODOS

A abrangência da análise se referiu ao Município do Rio de Janeiro e aos seus bairros (Anexo 1). Foram utilizadas as informações do Banco de Dados da Polícia Civil do Estado do Rio de Janeiro, que contém os Registros de Ocorrência (RO), da Secretaria de Estado de Polícia Civil do Rio de Janeiro, no período de janeiro de 1998 a dezembro de 2000.

O Registro de Ocorrência consiste no documento que dá origem à abertura do processo e onde são registradas, na Polícia Civil, as primeiras informações sobre um crime. Uma vez que os acidentes de trânsito podem implicar em crimes contra a vida (homicídio culposo ou lesões corporais culposas) e contra o patrimônio, há a necessidade de se instaurar um inquérito e de se reunirem provas para a instrução do processo que levará a sentença. Verifica-se, então, que este documento apresenta informações mais relacionadas à gravidade do evento, seu local, data e hora de ocorrência, entre outras.

Tentou se identificar grupos com características semelhantes referentes aos bairros do Município do Rio de Janeiro. A formação dos grupos teve como base os tipos de acidentes de trânsito com vítimas (Lesão Corporal Culposa – Atropelamento; Lesão Corporal Culposa – Colisão; Homicídio Culposo – Atropelamento; Homicídio Culposo – Colisão).

Originalmente, o banco de dados constava de 217.670 registros distribuídos nos anos de 1998 (66.383), 1999 (88.368) e 2000 (62.919) em todo o Estado do Rio de Janeiro. Após o primeiro recorte, para o Município do Rio de Janeiro, restaram 110.913

registros distribuídos em 1998 (34.921), 1999 (45.967) e 2000 (30.025). Como o objeto de estudo é o evento “acidente de trânsito” e não a “vítima de acidente de trânsito”, esses registros foram selecionados para representar cada um dos Registros de Ocorrência, da Polícia Civil, independente do número de vítimas que estivessem envolvidas no evento. Este banco de dados era composto de 36.012 registros de ocorrências de acidentes de trânsito com vítimas ocorridos no Município do Rio de Janeiro, no período de 01/01/1998 a 31/12/2000. Porém, em alguns registros não foi possível identificar de forma clara o bairro de ocorrência do evento. Foi realizada então uma análise de cada um dos registros, para verificar se o bairro de ocorrência estava classificado corretamente. Utilizou-se o Guia de Ruas Páginas Amarelas (1992) e o Guia Quatro Rodas (2002) para classificar os bairros de acordo com as ruas de ocorrência. Assim, para cada um dos registros foi verificada se a informação do bairro de ocorrência do evento fornecida pela Polícia Civil correspondia ao bairro do endereço registrado, quando essa informação estava disponível. Nos registros onde não constava de forma clara o endereço da ocorrência do acidente, ou quando o endereço não estava preenchido, foi considerado o bairro registrado pela Polícia Civil no banco de dados.

Assim, após a análise realizada com o objetivo de depurar as informações constantes no banco de dados original e também corrigir os possíveis erros de classificação e de digitação da variável bairro de ocorrência, restaram 35.278 (98,0%) registros de acidentes de trânsito envolvendo vítimas (fatais e não fatais) de acidentes de trânsito, no Município do Rio de Janeiro, para o triênio 1998-2000.

Os registros de acidentes de trânsito com vítimas foram representados segundo ano de ocorrência do evento, tipo de evento, dia de ocorrência na semana e hora de ocorrência.

Para a análise espacial exploratória, inicialmente foram construídos mapas cloropléticos para a visualização do número de ocorrências de acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas, para o triênio 1998-2000, nos 157 bairros que formam o Município do Rio de Janeiro. Para a construção dos mapas cloropléticos, utilizou-se o software de domínio público CSPro, versão 2.6, desenvolvido e distribuído pelo *U.S. Census Bureau* e obtido através do site www.census.gov/ipc/www/cspro. A análise exploratória se deu também através da análise das propriedades de primeira ordem (tendência) e de segunda ordem (dependência espacial).

A malha digital com a representação geográfica dos bairros do Município do Rio de Janeiro foi obtida através do Centro de Informação Científica e Tecnológica (CICT), da Fundação Oswaldo Cruz.

A análise das propriedades de primeira ordem baseou-se na Média Móvel Espacial. Foi criada a matriz de vizinhanças para utilização nas análises espaciais e adotou-se como critério de vizinhança a conectividade entre áreas. Assim, se o bairro *j* era vizinho (adjacente) do bairro *i*, o mesmo recebeu valor 1 na matriz de vizinhanças, caso contrário recebeu valor 0. A matriz de vizinhanças foi normalizada, ou seja, a soma dos pesos de cada uma das linhas é igual a um. Para analisar as propriedades de segunda ordem que estão relacionadas à dependência espacial, foram utilizados o I de Moran, o C de Geary e o diagrama de espalhamento de Moran. Na avaliação dos coeficientes de correlação espacial, I de Moran e C de Geary, foi utilizado o teste de independência espacial com um nível de significância de 5%.

Para a análise espacial exploratória utilizou-se o software de domínio público R, versão 2.0.1, desenvolvido por John Chambers e colaboradores, através da *Bell Laboratories* (antiga AT&T e atualmente *Lucent Technologies*) e a biblioteca *spdep* (*Spatial dependence: weighting schemes, statistics and models*) desenvolvida por Roger Bivand, que foram obtidos através do site www.r-project.org.

A escolha das variáveis explicativas utilizadas na construção dos modelos estudados foi baseada na experiência relatada em outros estudos, já descritos anteriormente, e está diretamente relacionada a sua obtenção junto às fontes oficiais de informação. São elas:

- a) **pop** - população residente na área (bairro) em estudo em 2000;
- b) **area** - área do bairro em Km²;
- c) **idh** - Índice de Desenvolvimento Humano da área (bairro) em 2000;
- d) **sanea** - percentual de domicílios com saneamento básico na área (bairro) em estudo em 2000;
- e) **menor18** - proporção da população residente na área (bairro) em estudo em 2000, com idade inferior a 18 anos;
- f) **maior60** - proporção da população residente na área (bairro) em estudo em 2000, com idade superior a 60 anos;
- g) **homem** - proporção de homens na população residente na área (bairro) em estudo em 2000;
- h) **casado** - proporção pessoas casadas na população residente na área (bairro) em estudo em 2000;

- i) **superior** - proporção de pessoas com nível superior de escolaridade na população residente na área (bairro) em estudo em 2000;
- j) **renda** - renda mediana da área (bairro) em 2000;
- k) **velocmed** - velocidade média das vias inseridas em cada da área (bairro);
- l) **via_ap** - quantidade total de vias arteriais primárias inseridas em cada da área (bairro);
- m) **via_as** - quantidade total de vias arteriais secundárias inseridas em cada da área (bairro);
- n) **via_c** - quantidade total de vias coletoras inseridas em cada da área (bairro);
- o) **via_e** - quantidade total de vias expressas inseridas em cada da área (bairro);
- p) **via_l** - quantidade total de vias locais inseridas em cada da área (bairro);
- q) **lin_bus** - quantidade total de linhas de ônibus que trafegam em cada da área (bairro);
- r) **intersec** - quantidade de intersecções (junções) de vias em cada da área (bairro).

As variáveis pop, sanea, menor18, maior60, homem, casado e superior foram obtidas através das informações do Censo Demográfico 2000, da Fundação Instituto Brasileiro de Geografia e Estatística (IBGE).

A variável sanea foi construída a partir das variáveis “Forma de abastecimento de água” e “Escoadouro do banheiro ou sanitário”, do Censo Demográfico 2000, do IBGE. Foram considerados com saneamento básico os domicílios que responderam que a forma de abastecimento de água era a rede geral e cujo escoadouro do banheiro ou sanitário era rede geral de esgoto ou pluvial ou então fossa séptica.

As variáveis area, idh e renda foram obtidas através do Instituto Municipal de Urbanismo Pereira Passos (IPP).

As variáveis velocmed, via_ap, via_as, via_c, via_e, via_l, lin_bus e intersec foram cedidas pelo Laboratório Planet, do Programa de Engenharia de Transportes (PET), do Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia (COPPE), da Universidade Federal do Rio de Janeiro (UFRJ).

Segundo o Código de Trânsito Brasileiro (BRASIL, 1997), via é a superfície por onde transitam veículos, pessoas e animais, compreendendo a pista, a calçada, o acostamento, ilha (obstáculo físico, colocado na pista de rolamento, destinado à ordenação dos fluxos de trânsito em uma interseção) e canteiro central (obstáculo físico construído como separador de duas pistas de rolamento, eventualmente substituído por marcas viárias (canteiro fictício)). As vias de tráfego são classificadas em:

- a) Via Expressa: Vias destinadas a atender grandes volumes de tráfego de longa distância e de passagem e a ligar os sistemas viários urbano, metropolitano e regional, com elevado padrão de fluidez;
- b) Via Arterial I e II (Primária e Secundária): Vias destinadas a absorver substancial volume de tráfego de passagem de média e longa distância, a ligar pólos de atividades, a alimentar vias expressas e estações de transbordo e carga, conciliando estas funções com a de atender ao tráfego local, com bom padrão de fluidez;
- c) Via Coletora: Vias destinadas a coletar o tráfego das vias locais e distribuí-lo nas vias arteriais e expressas, a servir de rota de transporte coletivo e a atender na mesma proporção o tráfego de passagem e local com razoável padrão de fluidez;

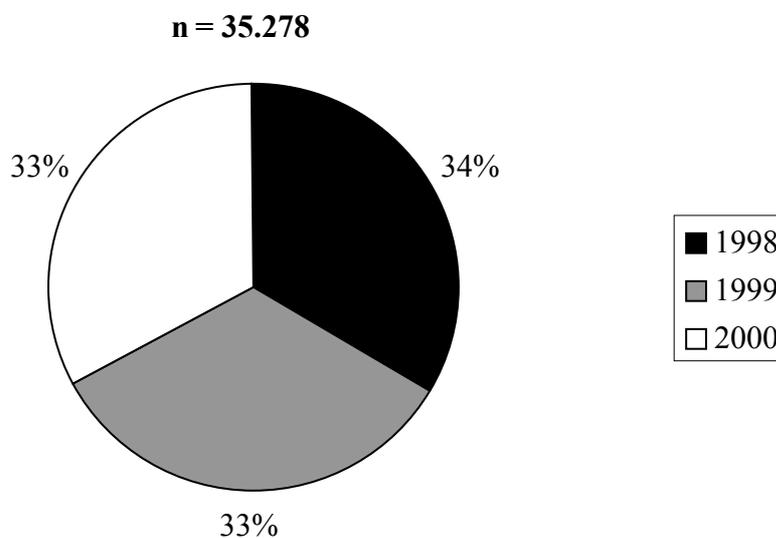
- d) Via Local: Vias destinadas a atender ao tráfego local, de uso predominantemente nesta via, com baixo padrão de fluidez.

5. RESULTADOS

5.1. Análise Exploratória

A análise exploratória das variáveis ano de ocorrência do evento, tipo de evento, dia de ocorrência na semana e hora de ocorrência dos registros de acidentes de trânsito com vítimas, no Município do Rio de Janeiro, para o triênio 1998-2000, estão descritas a seguir.

A figura 2 apresenta a distribuição dos registros de acidentes de trânsito com vítimas, no Município do Rio de Janeiro, segundo ano de ocorrência. Observa-se que o número de registros é aproximadamente o mesmo para cada um dos anos estudados.



Fonte: Banco de dados da PCERJ

Figura 2: Distribuição dos registros segundo ano de ocorrência.

A classificação do registro segundo tipo de evento (Homicídio - Colisão; Homicídio - Atropelamento; Lesão Corporal Culposa - Atropelamento; Lesão Corporal Culposa - Colisão) pode ser observada na figura 3. Verifica-se que a maioria dos registros de acidentes de trânsito envolvendo vítimas se deve aos atropelamentos e que estes são os maiores responsáveis pelos óbitos ocorridos nesse tipo de agravo. Cabe ressaltar que um mesmo acidente pode ter sido classificado como mais de um tipo de evento. Por exemplo, em uma colisão com duas vítimas, se uma delas morrer e a outra não, tem-se um Homicídio Culposos - Colisão e uma Lesão Corporal Culposa – Colisão. Assim, a figura a seguir apresenta todos os tipos de evento registrados nos 35.278 casos notificados. Do total de casos notificados no período estudado, 269 registros tiveram duas classificações e 3 registros tiveram três tipos de evento relacionados.

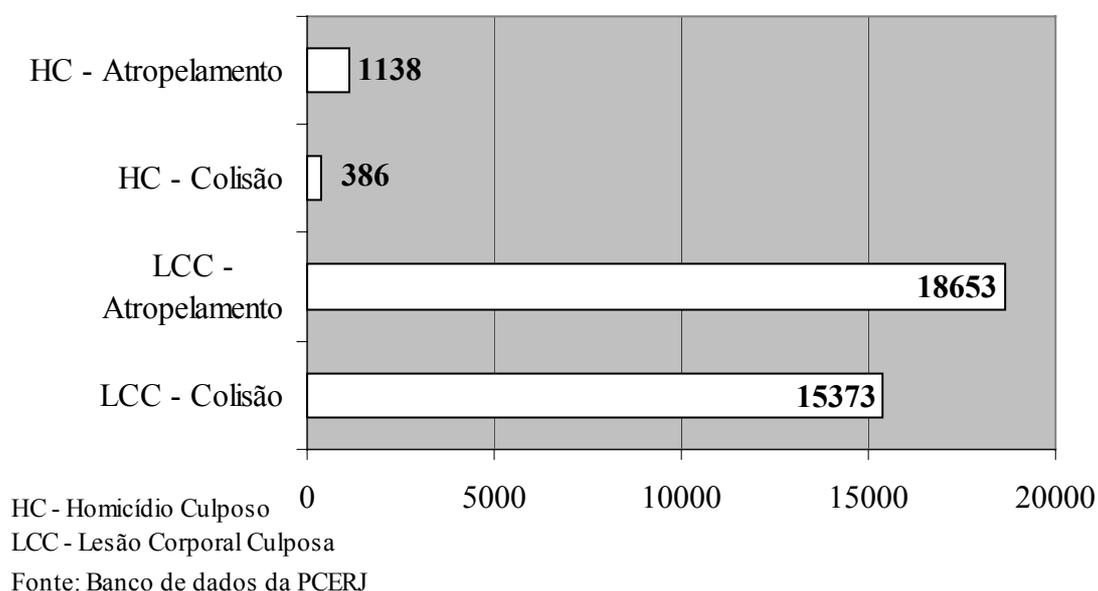
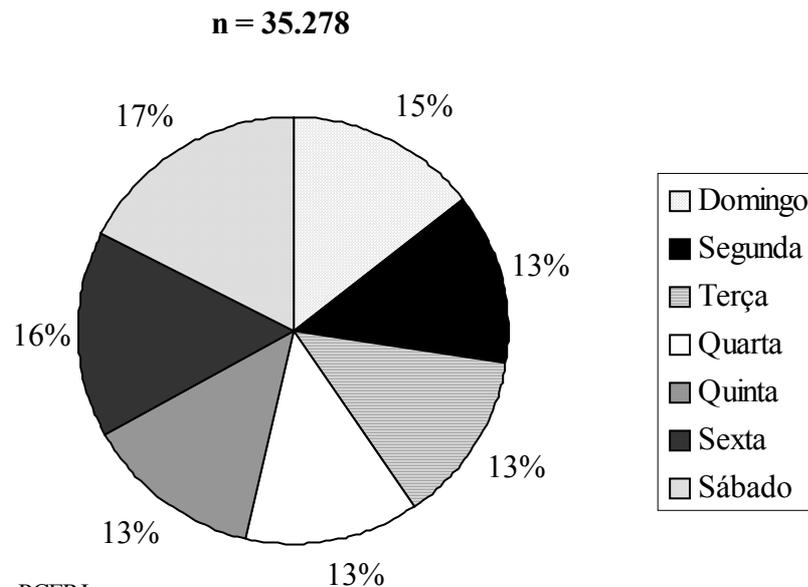


Figura 3: Distribuição dos registros segundo tipo de evento.

Como pode ser visto na figura 4, os finais de semana, ou seja, a sexta-feira (16%), o sábado (17%) e o domingo (15%) foram responsáveis por 48% dos registros de ocorrências de acidentes de trânsito com vítimas, no Município do Rio de Janeiro, no período de 01/01/1998 a 31/12/2000.



Fonte: Banco de dados da PCERJ

Figura 4: Distribuição dos registros segundo dia da semana.

Em relação à hora de ocorrência do evento, verifica-se através da figura 5 que a maior incidência se deu no período das 12:01 as 18:00 horas (34%) seguido do período das 18:01 as 24:00 horas (31%). Este fato pode ser explicado pela hora do rush onde há um maior número de veículos em circulação assim como um maior número de pedestres nas ruas.

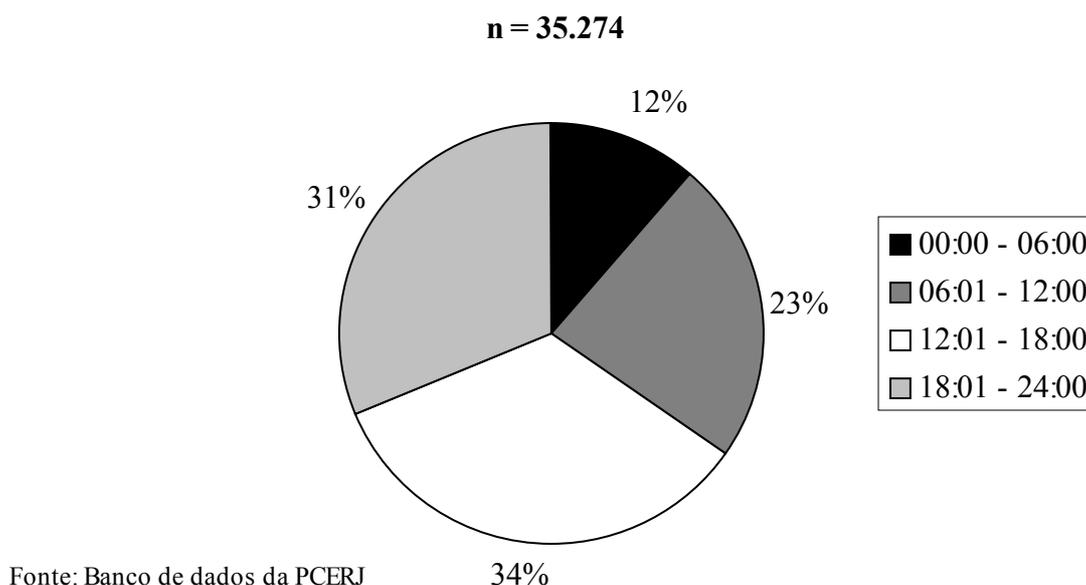


Figura 5: Distribuição dos registros segundo hora do evento.

Os dados foram agrupados segundo bairro de ocorrência para se construir a distribuição espacial do evento em estudo. A distribuição do número total de acidentes de trânsito com vítimas no Município do Rio de Janeiro, no período de 1998 a 2000, segundo ano de ocorrência e bairro de ocorrência encontra-se representada no Anexo 2. Segundo o tipo de evento, os acidentes de trânsito com vítimas foram classificados em colisões com vítimas e atropelamentos, e a distribuição segundo tipo de evento encontra-se no Anexo 3.

5.2. Análise Espacial

A figura 6 apresenta o boxplot para a distribuição do número de ocorrências de acidentes de trânsito com vítimas, de colisões com vítimas e de atropelamentos, registrados nos bairros do Município do Rio de Janeiro, no período de 1998 a 2000. Através dessa figura pode-se verificar que os três tipos de eventos estudados apresentaram pontos discrepantes (outliers) na sua distribuição dos dados, ou seja, pontos acima de 1,5 vezes o desvio interquartil mais o terceiro quartil. Foram 16 outliers na distribuição dos acidentes de trânsito com vítimas, 17 outliers na distribuição das colisões e 14 outliers na distribuição dos atropelamentos. Os bairros de Campo Grande, Centro, Barra da Tijuca, Bangu, São Cristóvão, Realengo, Santa Cruz, Bonsucesso, Jacarepaguá, Tijuca, Copacabana, Madureira e Botafogo foram considerados discrepantes nos três tipos de eventos relacionados. Além desses, Penha e Taquara apresentaram valores altos nos acidentes de trânsito com vítimas e nas colisões com vítimas. O Méier apresentou valores altos nos acidentes de trânsito com vítimas e nos atropelamentos. Leblon e Irajá foram bairros atípicos somente na distribuição das colisões com vítimas.

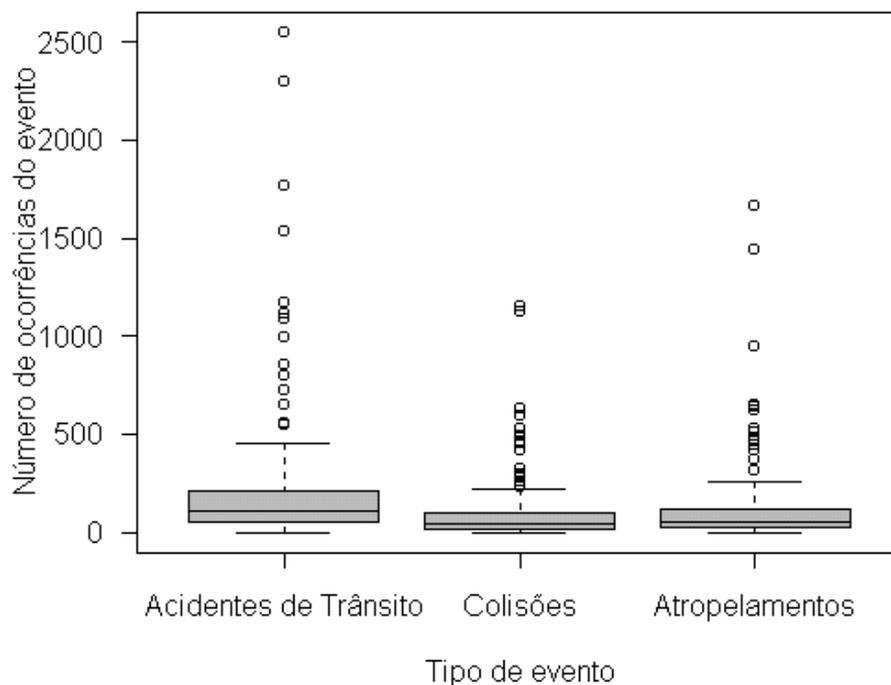


Figura 6: Boxplot para a distribuição do número de ocorrências de acidentes de trânsito com vítimas, de colisões com vítimas e de atropelamentos.

5.2.1. Análise Espacial dos Acidentes de Trânsito com Vítimas

A distribuição espacial dos registros de acidentes de trânsito com vítimas, no triênio 1998-2000 pode ser observada na figura 7. Nesta figura adotou-se a divisão das classes segundo quintis. A opção por quintis para a divisão das classes ocorreu uma vez que, por ser uma separatriz, essa medida de posição não sofre influência de valores extremos da série e além disso na divisão por classes de intervalos de mesma amplitude optou-se por cinco intervalos. Assim, dos 157 bairros que compõem o Município do Rio de Janeiro, os 31 (20%) bairros assinalados com o tom mais claro de cinza são aqueles

que apresentaram o menor número de casos registrados no período estudado, variando de um único caso, registrado na Praia da Bandeira, na Ilha do Governador, até 40 casos registrados nos bairros do Jacarezinho, Mangueira e Rocha. Da mesma forma, os 33 bairros assinalados com o tom mais escuro de cinza são aqueles que apresentaram o maior número de casos registrados no período, variando de 239 casos, registrados nos bairros de Guadalupe e Vicente de Carvalho até 2.546 casos no bairro de Campo Grande.

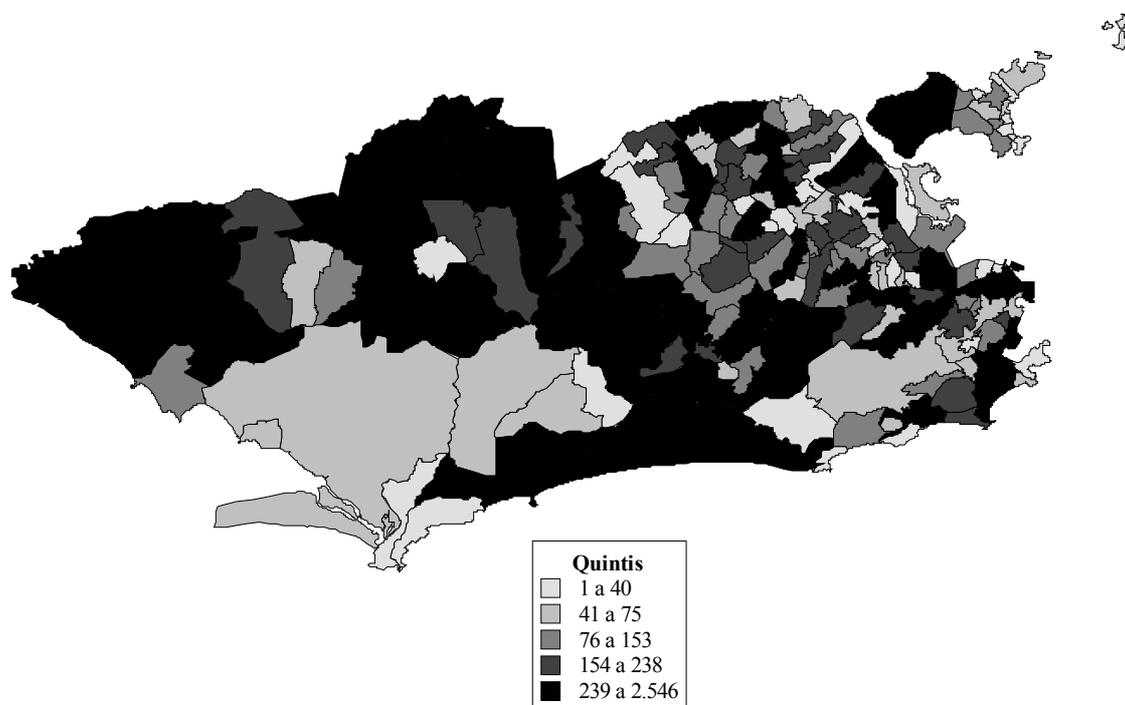


Figura 7: Distribuição por quintis dos acidentes de trânsito com vítimas.

Triênio 1998-2000.

A figura 8 apresenta a distribuição espacial dos registros de acidentes de trânsito com vítimas, no triênio 1998-2000 utilizando como critério de divisão das classes a mesma amplitude. Utilizando esse critério observa-se, mais claramente, os bairros com

as maiores incidências do evento. No último intervalo de classes (2.037 – 2.546) encontram-se os bairros de Campo Grande (2.546 registros) e Centro (2.296 registros). No intervalo de 1.528 até 2.036 registros encontram-se os bairros da Barra da Tijuca (1.770 registros) e Bangu (1.533 registros). No intervalo de 1.019 até 1.527 registros encontram-se os bairros de São Cristóvão (1.175 registros), Realengo (1.112 registros) e Santa Cruz (1.088 registros). Com a análise dessa figura chama a atenção a grande concentração de acidentes em bairros localizados na Zona Oeste do Município e também na região Central da Cidade. Neste sete bairros juntos ocorreram 11.520 acidentes com vítimas que representam 32,7% do total de acidentes de trânsito com vítimas registrado para o Município do Rio de Janeiro no período de 01/01/1998 a 31/12/2000.

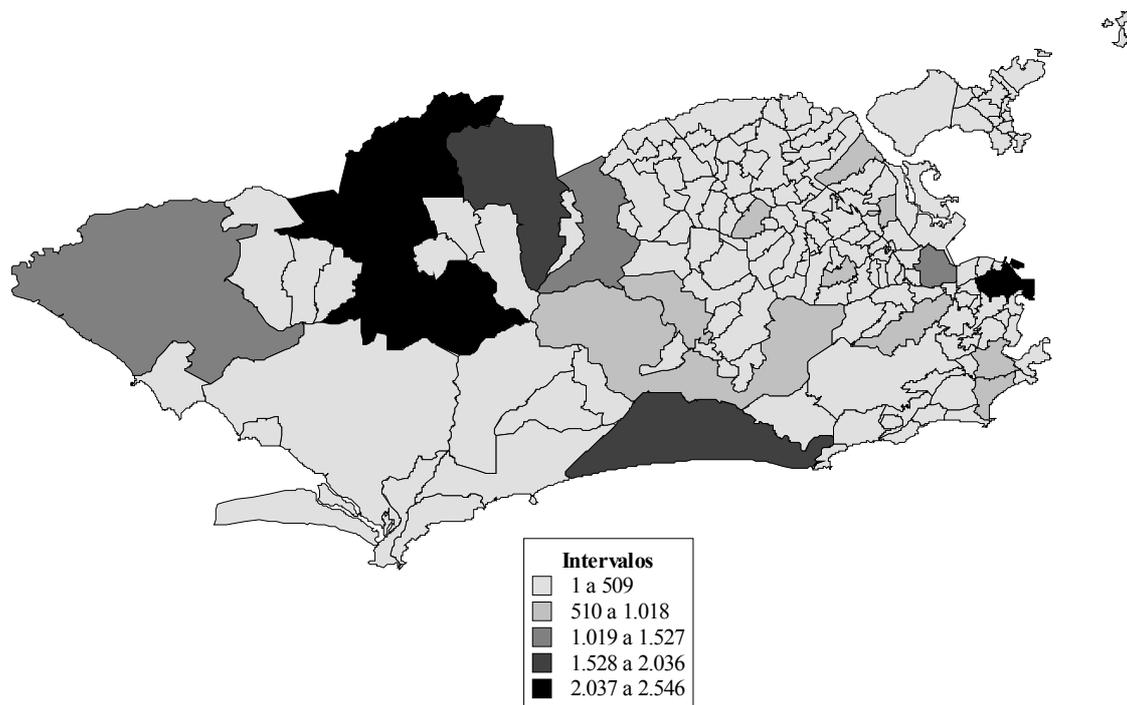


Figura 8: Distribuição por intervalos dos acidentes de trânsito com vítimas.

Triênio 1998-2000.

A distribuição espacial das médias móveis espaciais dos registros de acidentes de trânsito com vítimas, no triênio 1998-2000 podem ser observadas nas figuras 9 e 10. Na figura 9 foi utilizada a divisão das classes segundo quintis e na figura 10 foram de faixas de mesma amplitude.

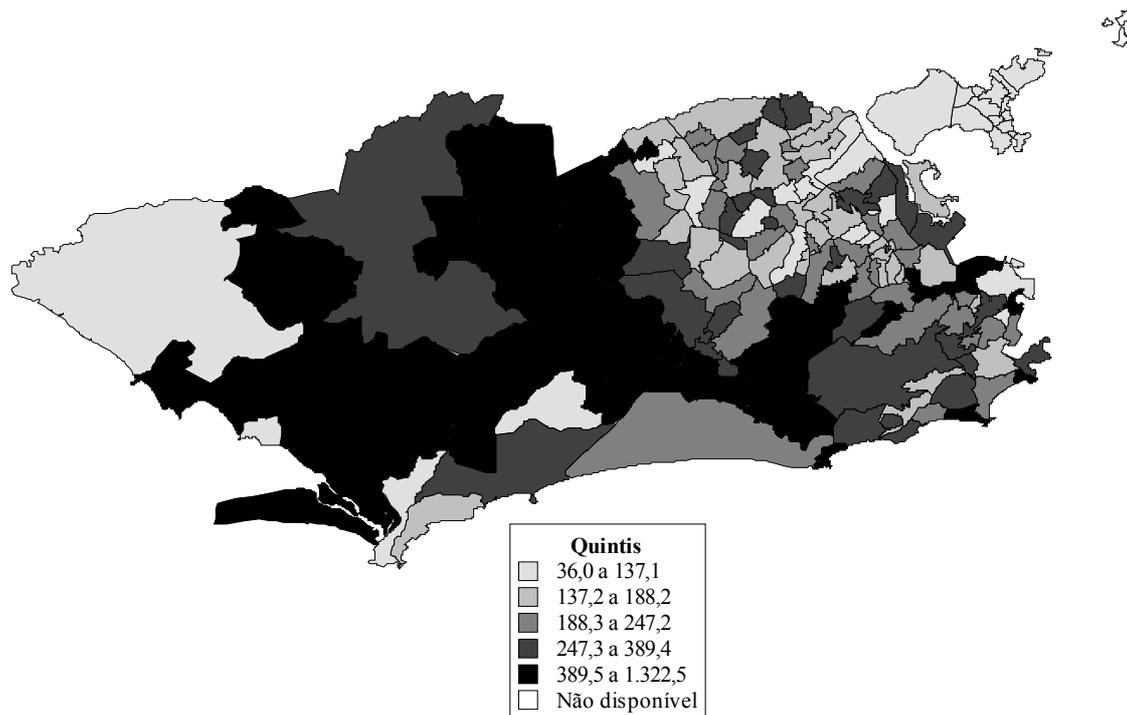


Figura 9: Médias móveis por quintis dos acidentes de trânsito com vítimas.

Triênio 1998-2000.

Observa-se que a aplicação desta técnica de suavização dos dados provocou uma grande diminuição no número máximo de casos de uma área que antes era no bairro de Campo Grande que registrou 2.546 casos e que após a suavização passou a ser o bairro de Padre Miguel com 1.322,5 registros. A partir de Padre Miguel os bairros que registraram um maior número de casos foram Paciência (1.236 registros), Saúde (1.158,5 registros) e Santíssimo (1.072,2 registros). Destaca-se que os bairros de

Paciência e Santíssimo são adjacentes ao bairro de Campo Grande enquanto o bairro da Saúde faz fronteira com o Centro. O bairro de Padre Miguel só possui fronteiras com os bairros de Realengo e Bangu e teve uma média móvel espacial tão elevada, em comparação com as demais localidades estudadas, uma vez que seus vizinhos tiveram uma grande concentração de ocorrências. Destaca-se ainda que com a aplicação das médias móveis espaciais o bairro do Centro passou a figurar entre aqueles com os menores índices de ocorrência.

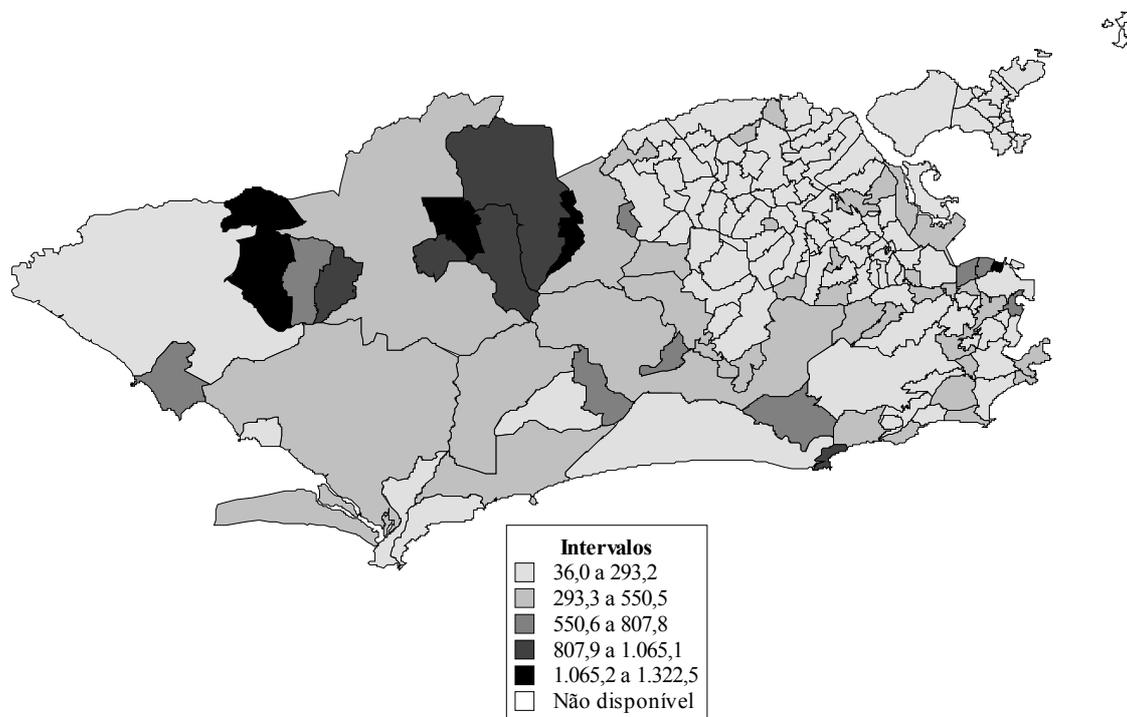


Figura 10: Médias móveis por intervalos dos acidentes de trânsito com vítimas.

Triênio 1998-2000.

Para verificar a estrutura de autocorrelação espacial foram utilizados o I de Moran e o C de Geary, ambos com a vizinhança por conectividade de áreas, para o total de eventos no triênio 1998-2000. Seus resultados foram, respectivamente, -0,0355 e

1,2784. Os testes de independência espacial para o I de Moran e o C de Geary demonstraram não haver correlação espacial entre os dados, uma vez que a hipótese nula (H_0), ou seja, a hipótese de completa aleatoriedade espacial, não foi rejeitada, resultando, respectivamente, nos seguintes valores probabilísticos: $p=0,7255$ e $p=0,9970$.

A baixa correlação espacial pode ser visualizada pelo diagrama de espalhamento de Moran (figura 11), que apresentou a maioria dos pontos localizados no segundo e quarto quadrantes, confirmando assim, os resultados encontrados nos testes de independência espacial.

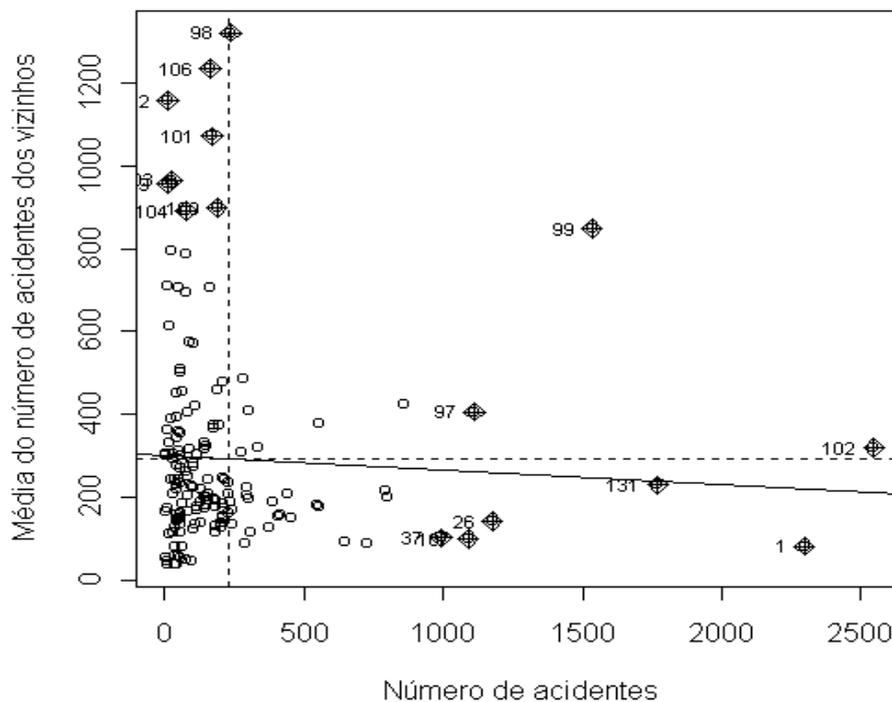


Figura 11: Diagrama de espalhamento de Moran para os acidentes de trânsito com vítimas. Triênio 1998-2000.

5.2.2. Análise Espacial dos Atropelamentos

Os registros de atropelamentos ocorridos, no triênio 1998-2000 podem ser observados na figura 12, que utilizou os quintis como critério de divisão de classes. Os bairros assinalados com o tom mais claro de cinza são aqueles que apresentaram o menor número de casos registrados no período, variando de nenhum caso, registrados na Praia da Bandeira e no Campo dos Afonsos, até 21 casos registrados nos bairros Bancários, localizado na Ilha do Governador, e Rocha. Os bairros assinalados com o tom mais escuro de cinza foram os que apresentaram o maior número de casos registrados no período, variando de 142 casos, registrados no bairro de Ramos até 1.665 casos no registrados no Centro.

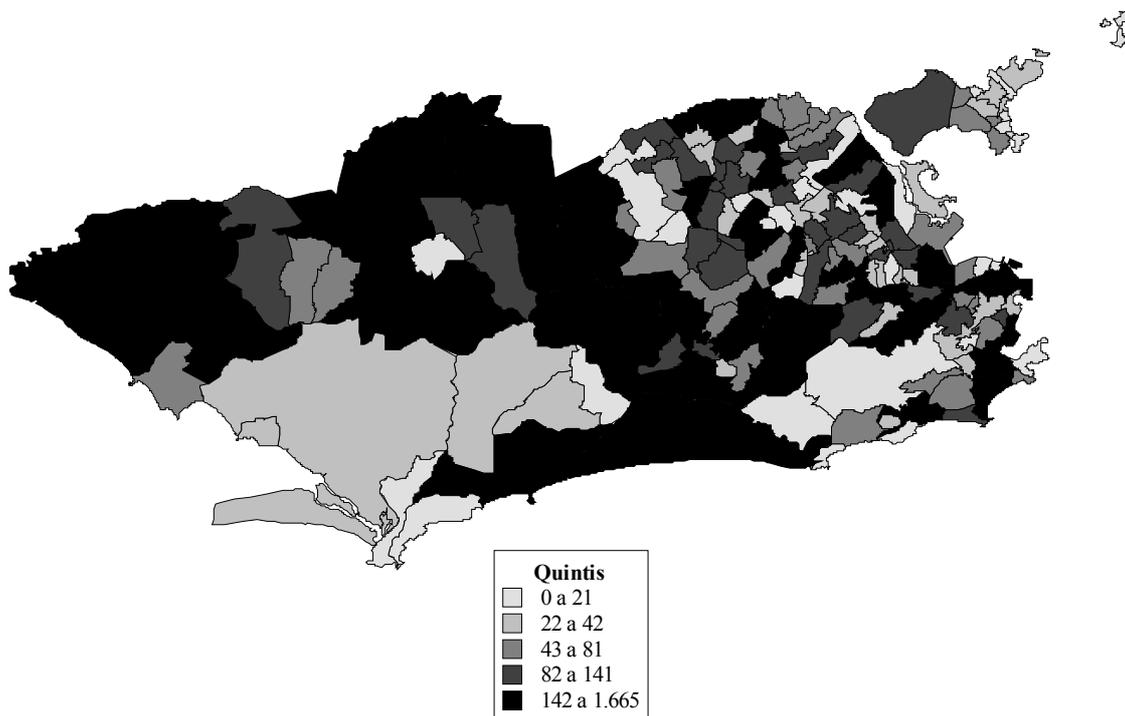


Figura 12: Distribuição por quintis dos atropelamentos. Triênio 1998-2000.

A distribuição espacial dos registros de atropelamentos, no triênio 1998-2000 utilizando como critério de divisão das classes a mesma amplitude, pode ser observada na figura 13. Observa-se que os bairros com as maiores incidências do evento foram o Centro (1.665), Campo Grande (1.437) e Bangu (944). Destaca-se que nestes três bairros foram registrados 20,5% de todos os atropelamentos no período.

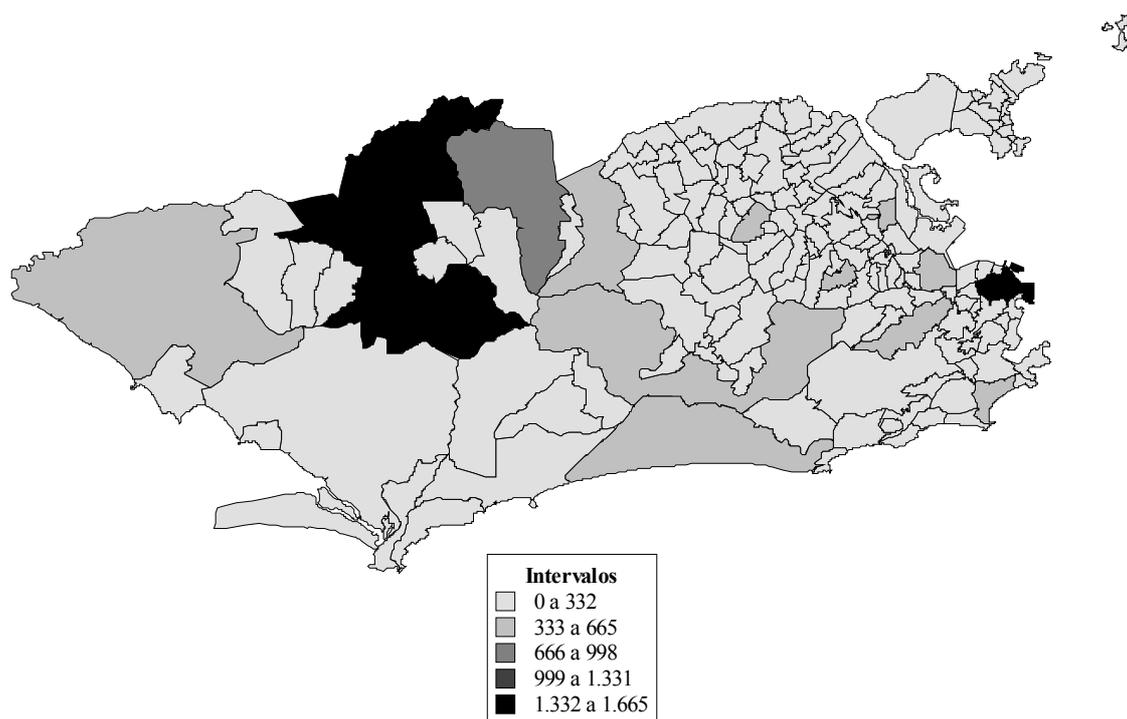


Figura 13: Distribuição por intervalos dos atropelamentos. Triênio 1998-2000.

A distribuição espacial das médias móveis espaciais dos registros de atropelamentos, no triênio 1998-2000 pode ser observada na figura 14, com a divisão das classes segundo quintis e na figura 15 com a divisão das classes através da mesma amplitude.

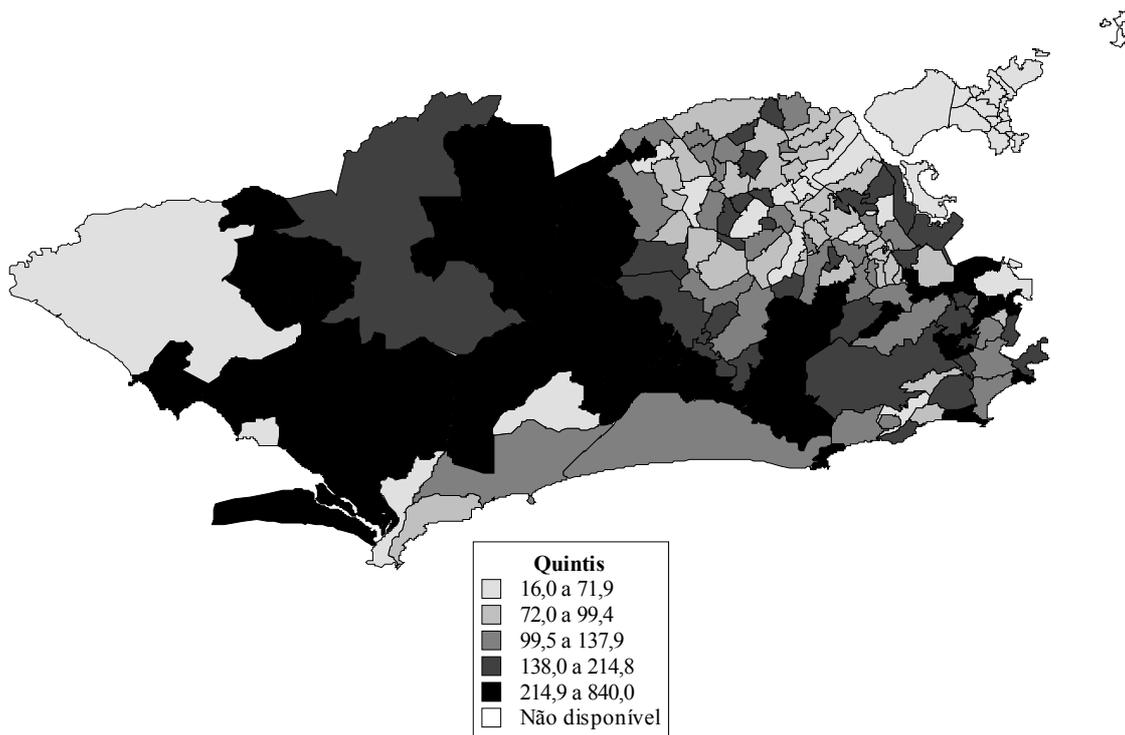


Figura 14: Médias móveis por quintis atropelamentos. Triênio 1998-2000.

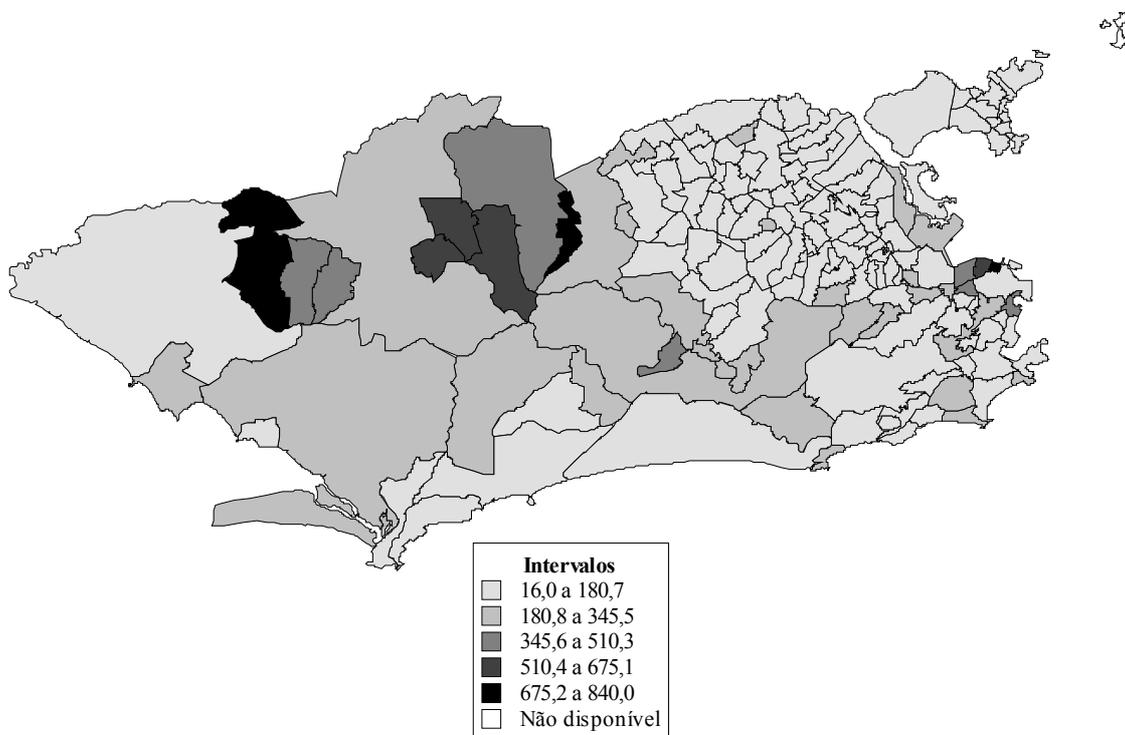
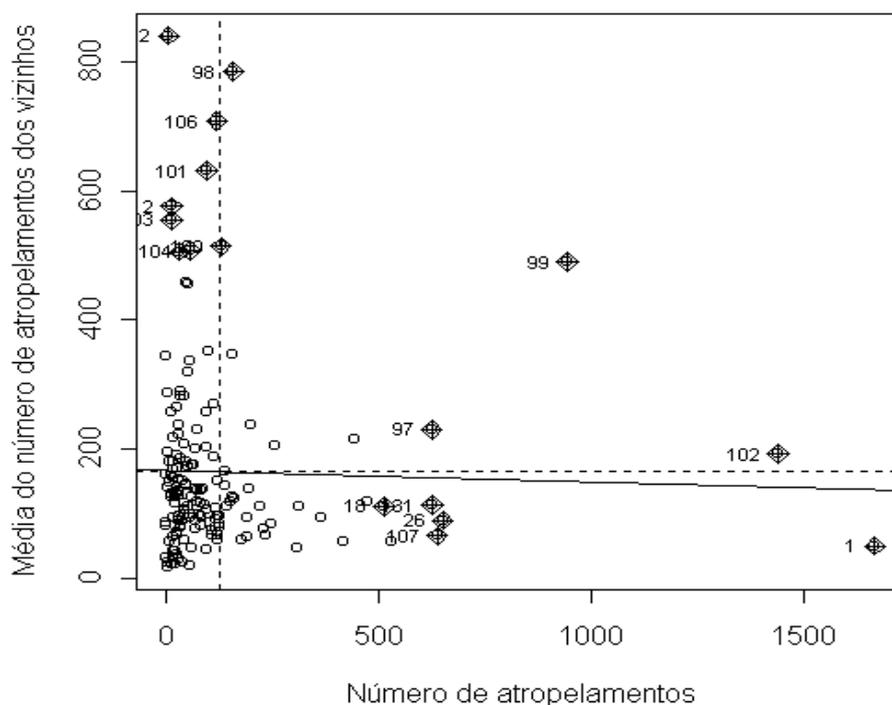


Figura 15: Médias móveis por intervalos dos atropelamentos. Triênio 1998-2000.

A suavização dos dados, obtida com a aplicação das médias móveis espaciais, provocou uma grande diminuição no número máximo de casos registrados. Assim, os bairros da Saúde (840), Padre Miguel (784,5), Paciência (707,7), e Santíssimo (631,2) foram os bairros que registraram um maior número de casos.

O I de Moran e o C de Geary, ambos com a vizinhança por conectividade, para o total de atropelamentos no triênio 1998-2000, tiveram como resultados, respectivamente, -0,0174 e 1,2619. Os testes de independência espacial para o I de Moran e para o C de Geary, também apontaram para uma completa aleatoriedade espacial entre os dados, indicando, assim, não haver correlação espacial entre as áreas. Os p-valores foram, respectivamente, 0,5914 e 0,9898.

O diagrama de espalhamento de Moran, apresentado na figura 16, corrobora os resultados encontrados, uma vez que a maioria dos pontos se localizaram no segundo e no quarto quadrantes, indicando então a falta de associação espacial.



**Figura 16: Diagrama de espalhamento de Moran para os atropelamentos.
Triênio 1998-2000.**

5.2.3. Análise Espacial das Colisões com Vítimas

A distribuição dos registros de colisão com vítimas, no triênio 1998-2000, utilizando a divisão de classes por quintis, pode ser observada através da figura 17. O Complexo do Alemão e a Maré não registraram nenhuma colisão com vítimas no período estudado. Barra da Tijuca, Campo Grande e Centro foram os bairros que apresentaram um maior número de registros no período, sendo, respectivamente, 1.149, 1.121 e 635 casos.

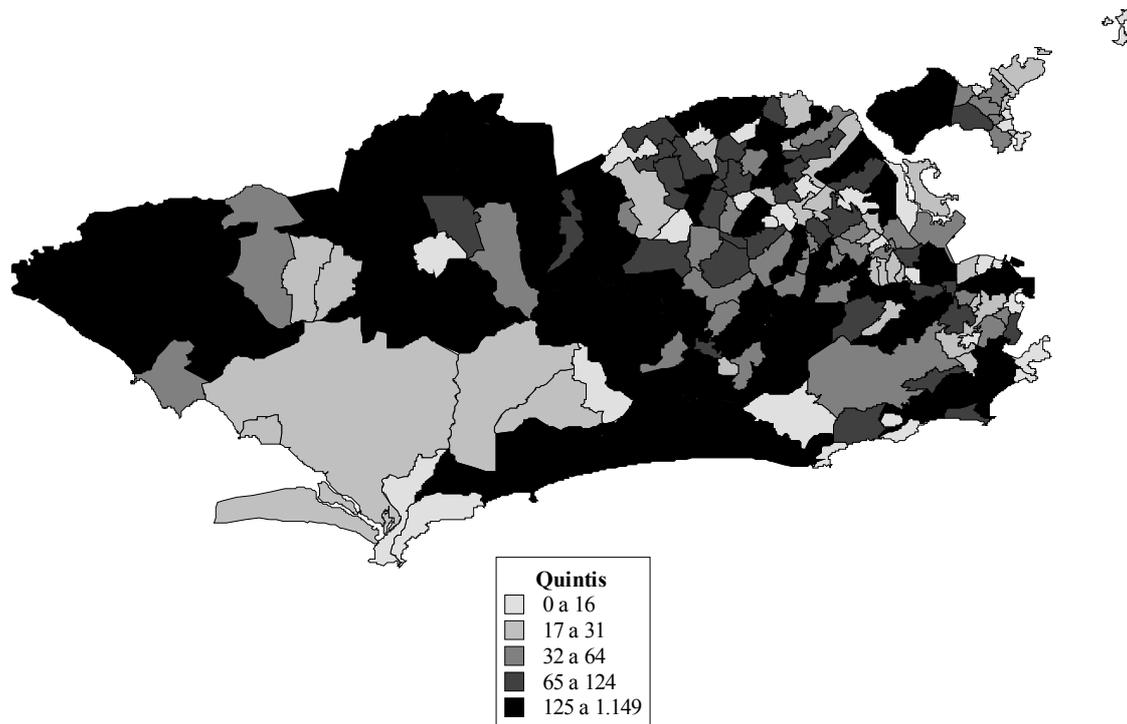


Figura 17: Distribuição por quintis das colisões com vítimas. Triênio 1998-2000.

A figura 18 apresenta a distribuição espacial dos registros de colisão com vítimas, no triênio 1998-2000. Foi utilizado como critério de divisão das classes a mesma amplitude. Barra da Tijuca e Campo Grande foram os bairros que apresentaram os maiores números de colisões com vítimas no período estudado. Deve-se ressaltar que essas duas localidades representaram 14,5% do total dos registros de colisão com vítimas.

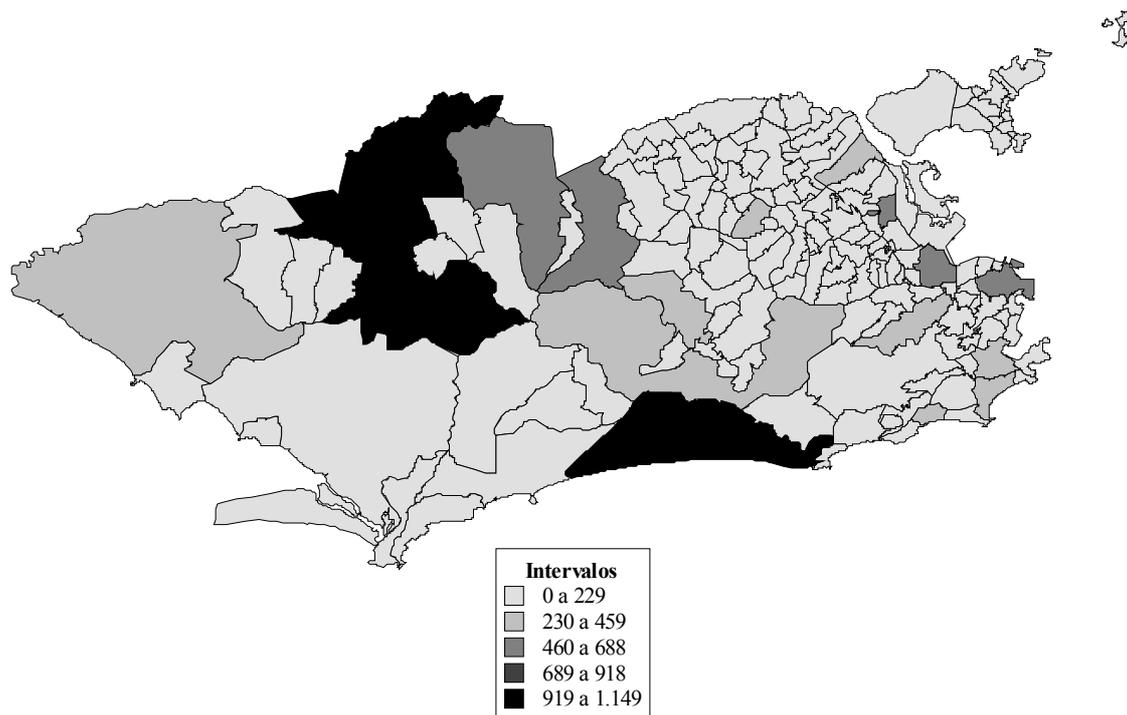


Figura 18: Distribuição por intervalos das colisões com vítimas. Triênio 1998-2000.

As médias móveis espaciais dos registros de colisão com vítimas, no triênio 1998-2000 estão representadas na figura 19 através de quintis e na figura 20 através de classes de mesma amplitude.

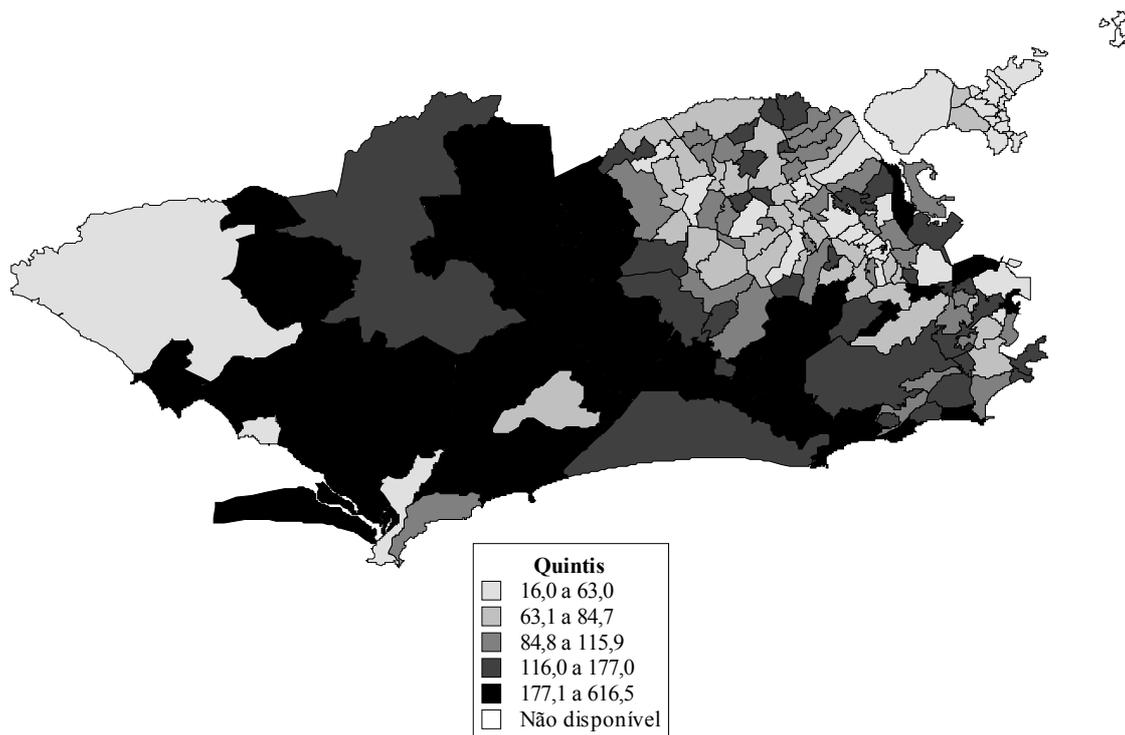


Figura 19: Médias móveis por quintis das colisões com vítimas. Triênio 1998-2000.

Através da suavização dos dados obtida pela aplicação das médias móveis espaciais houve uma grande redução no número de casos. Assim, a Barra da Tijuca que foi o bairro com o maior número de registros de colisão com vítimas no período, após a suavização passou a ter somente 117 casos. Em contrapartida, o seu vizinho, o bairro do Joá passou a ter o maior número de registros (616,5). Da mesma forma que bairros anteriormente sem registros de colisões passaram a ter, após a aplicação dessa técnica, casos registrados.

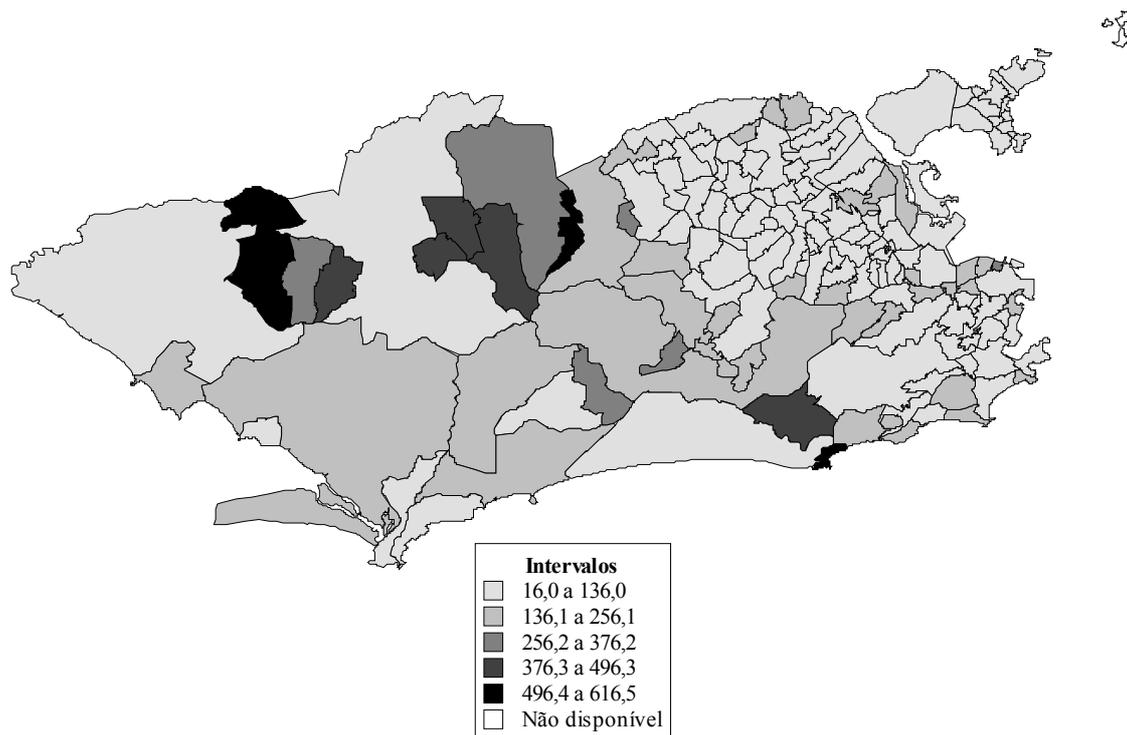


Figura 20: Médias móveis por intervalos das colisões com vítimas.

Triênio 1998-2000.

Os testes de correlação espacial para o I de Moran e para o C de Geary demonstraram não haver associação entre os dados. O I de Moran foi $-0,0452$ ($p=0,7893$) e o C de Geary, $1,2756$ ($p=0,9957$).

A figura 21 apresenta o diagrama de espalhamento de Moran e pode-se observar a baixa associação entre os dados.

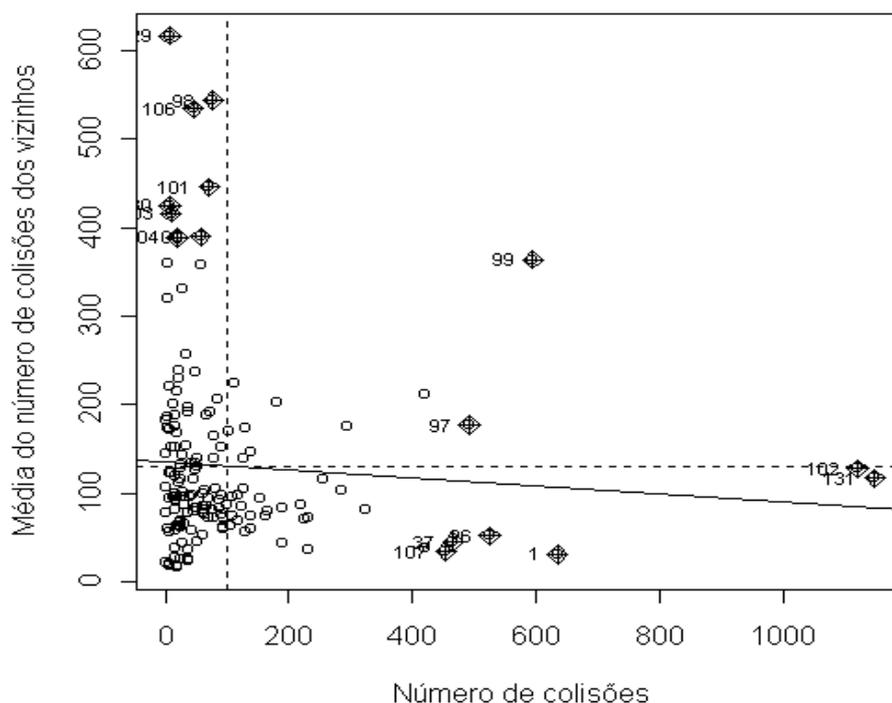


Figura 21: Diagrama de espalhamento de Moran para as colisões com vítimas.

Triênio 1998-2000.

Através da comparação da distribuição dos acidentes de trânsito com vítimas, dos atropelamentos e das colisões com vítimas observou-se que a maior concentração de casos ocorreu nos bairros do Centro da cidade e na Zona Oeste. Embora o bairro de Campo Grande tenha apresentado o maior número de casos de acidentes de trânsito com vítimas (2.546), verificou-se que nos atropelamentos o mesmo foi superado pelo Centro que registrou 1.665, contra 1.437. Em relação às colisões com vítimas, Campo Grande também ficou com o segundo maior número de casos (1.121), tendo sido superado pela Barra da Tijuca, que apresentou 1.149 registros.

Destaca-se também que enquanto o Centro e Campo Grande apresentaram, respectivamente, 2,62 e 1,28 registros de atropelamento para cada registro de colisão, na Barra da Tijuca essa relação ocorreu de forma inversa, ou seja, para cada registro de atropelamento houveram 1,84 registros de colisão.

5.3. Modelagem dos Dados

As estatísticas descritivas das variáveis explicativas selecionadas para o processo de modelagem encontram-se no quadro 6.

Quadro 6: Estatísticas descritivas (mínimo, máximo, quartil 1, mediana, quartil 3, média e desvio padrão) das variáveis explicativas utilizadas na modelagem.

Variáveis	Mínimo	Máximo	Q1	Mediana	Q3	Média	Desvio Padrão
pop	257	297500	12350	24850	46330	37310	42327,7
area	0,1611	139,5	1,375	2,618	4,866	7,742	19,3381
idh	0,7114	0,9704	0,7984	0,8389	0,878	0,8407	0,0656
sanea	0,2258	1	0,9155	0,9627	0,9896	0,9225	0,1209
menor18	0,1376	0,4637	0,2322	0,2706	0,3064	0,2713	0,0564
maior60	0,0396	0,2619	0,0925	0,1182	0,1525	0,1232	0,0448
homem	0,4184	0,5543	0,4584	0,4683	0,4803	0,4704	0,0219
casado	0,1504	0,5246	0,3258	0,3600	0,3823	0,3550	0,0526
superior	0,0113	0,7670	0,0948	0,1628	0,3051	0,2266	0,1762
renda	300	4250	480	600	1000	911,5	787,8
velocmed	15	44,02	18,13	20,47	23,09	21,1	4,4
via_ap	0	525	0	3	26	19,22	46,9
via_as	0	209	0	8	27	21,05	33,7
via_c	0	682	17	38	71	70,17	104,8
via_e	0	381	0	0	12	13,31	36,8
via_l	11	6798	78	163	347	385,3	756
lin_bus	0	574	43	72	118	105,5	100,6
intersec	16	5954	104	177	338	377,5	670,2

As figuras 22 a 30, apresentam os boxplots para as variáveis explicativas.

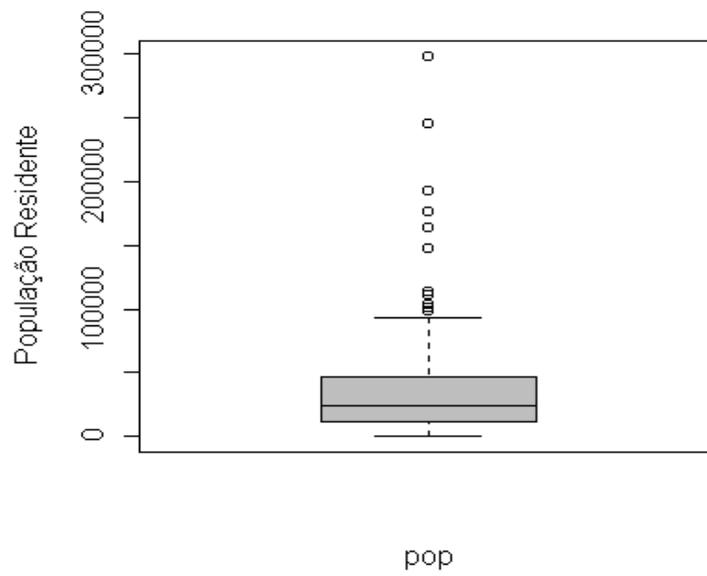


Figura 22: Boxplot para a distribuição da variável pop dos bairros do Município do Rio de Janeiro, em 2000.

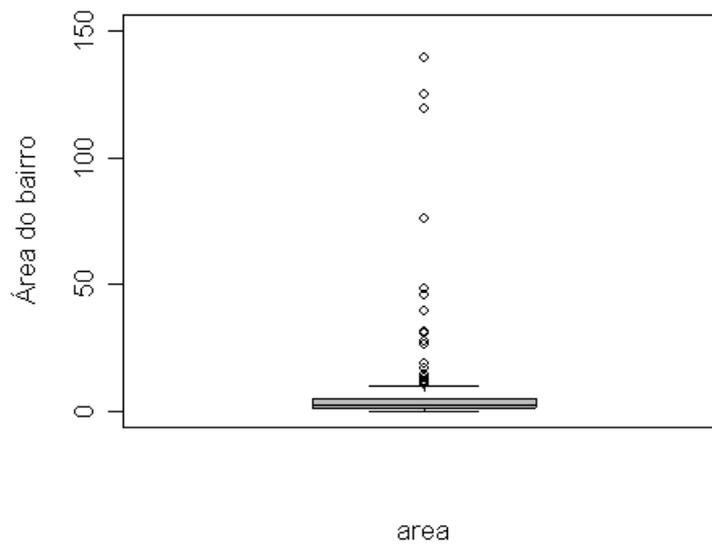


Figura 23: Boxplot para a distribuição da variável area dos bairros do Município do Rio de Janeiro.

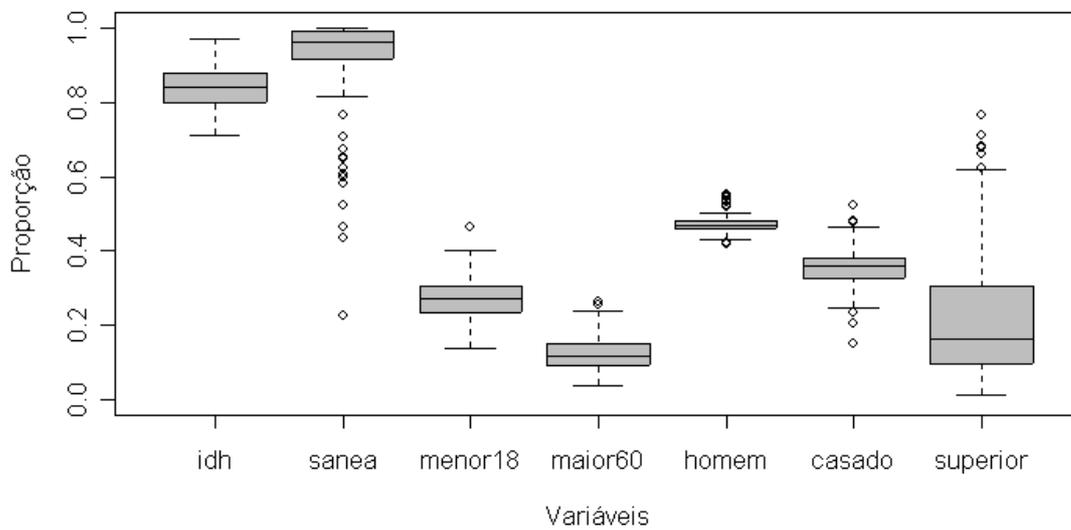


Figura 24: Boxplot para a distribuição das variáveis idh, sanea, menor18, maior60, homem, casado e superior dos bairros do Município do Rio de Janeiro, em 2000.

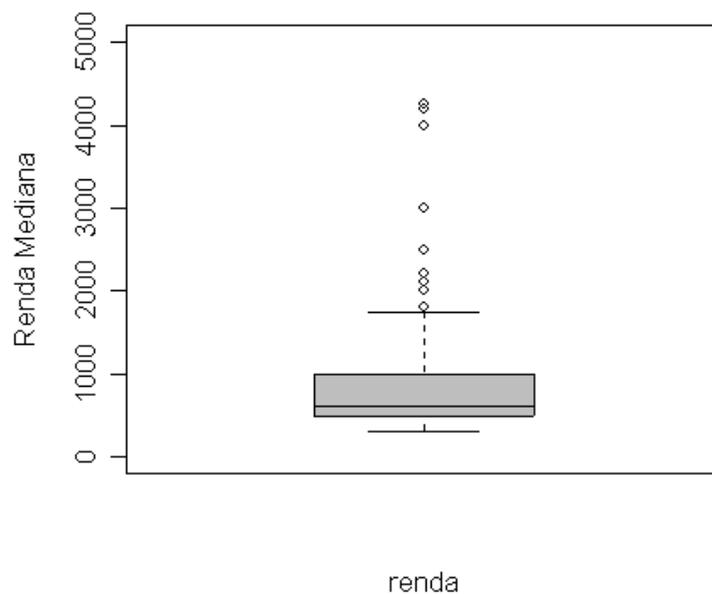


Figura 25: Boxplot para a distribuição da variável renda dos bairros do Município do Rio de Janeiro, em 2000.

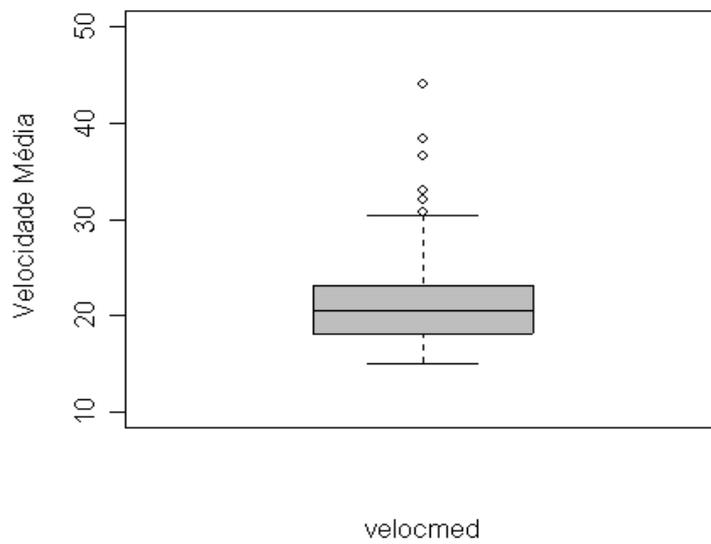


Figura 26: Boxplot para a distribuição da variável `velocmed` dos bairros do Município do Rio de Janeiro.

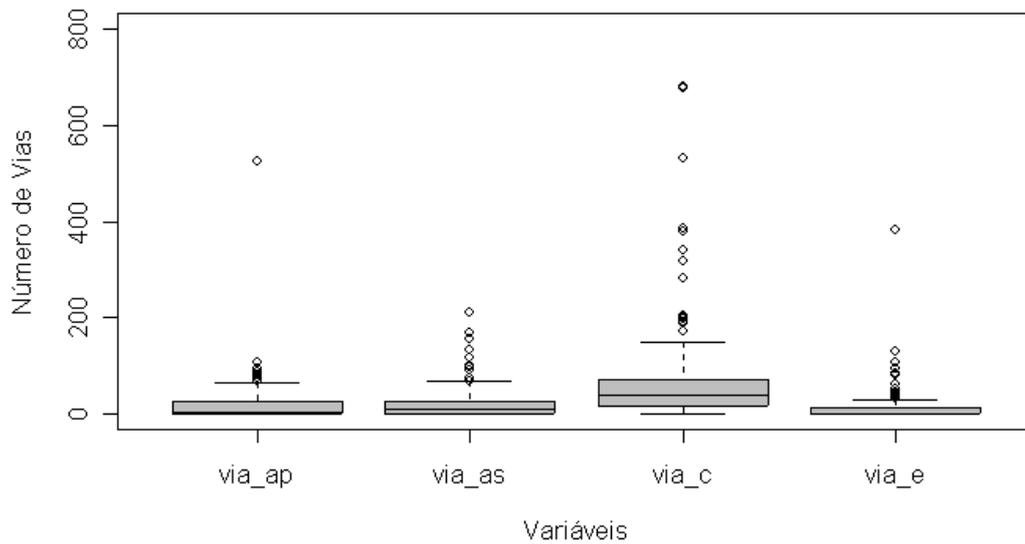


Figura 27: Boxplot para a distribuição das variáveis `via_ap`, `via_as`, `via_c` e `via_e` dos bairros do Município do Rio de Janeiro.

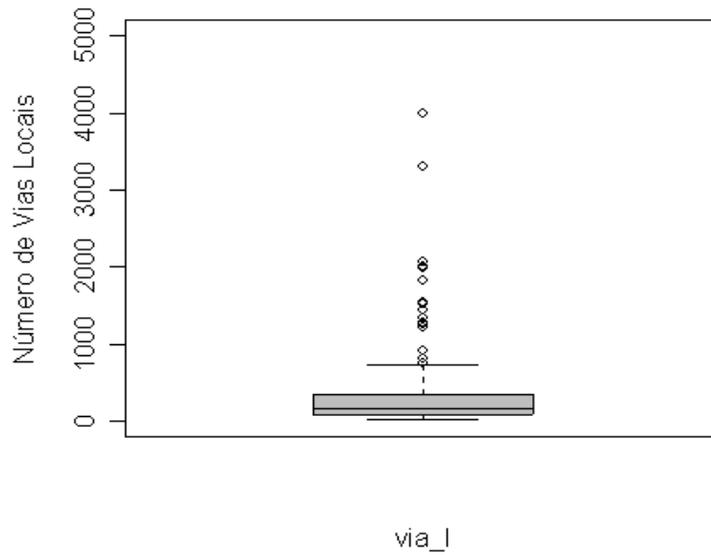


Figura 28: Boxplot para a distribuição da variável via_l dos bairros do Município do Rio de Janeiro.

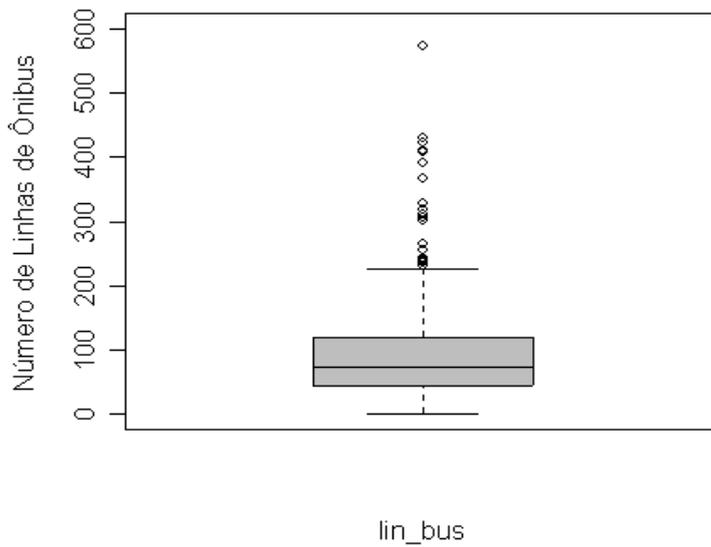


Figura 29: Boxplot para a distribuição da variável lin_bus dos bairros do Município do Rio de Janeiro.

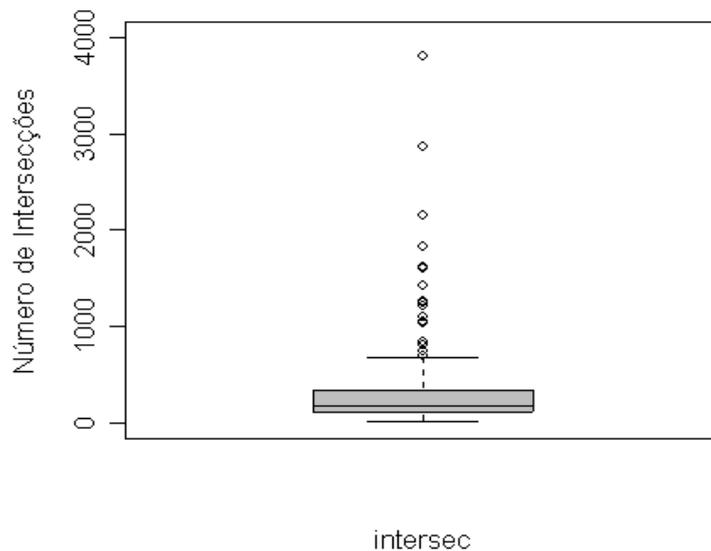


Figura 30: Boxplot para a distribuição da variável intersec dos bairros do Município do Rio de Janeiro.

A análise exploratória das variáveis explicativas realizada com base no quadro 6 e nas figuras 22 a 30, mostrou haver pouca variabilidade entre os bairros para as variáveis idh, maior60, homem. Estas variáveis apresentaram amplitude total (distância entre o maior e menor valor observados) de 0,2590, 0,2223 e 0,1359, além de desvio padrão de 0,0656, 0,0448 e 0,0219, respectivamente, indicando assim, haver pouca variabilidade nas suas respostas.

Sendo assim, para o processo de modelagem dos dados, serão utilizadas somente as variáveis sócio-demográficas pop, area, sanea, menor18, casado, superior e renda e as variáveis velocmed, via_ap, via_as, via_c, via_e, via_l, lin_bus e intersec que estão relacionadas ao trânsito.

O processo de modelagem consiste, basicamente na representação da realidade de uma forma mais simplificada. O processo envolve suposições à respeito da variável resposta e da sua relação com as variáveis explicativas. Como nesse estudo a variável resposta é uma contagem de eventos, a princípio os dados seguem uma Distribuição de Poisson. Contudo, para iniciar o processo de modelagem, primeiramente será analisada a variável resposta para identificar se realmente a mesma se distribui segundo uma Poisson.

Foi utilizado o Índice de Fisher para superdispersão (GÜRTLER e HENZE, 2000) para identificar a adequação dos dados a um Modelo de Poisson.

Foi verificado que os acidentes de trânsito com vítimas, as colisões com vítimas e os atropelamentos apresentavam superdispersão e, portanto, a contagem, de cada um desses tipos de eventos, não deveria ser modelada segundo uma Distribuição de Poisson.

Foi utilizada, então a Distribuição Binomial Negativa, uma vez que essa distribuição de probabilidades também é indicada para descrever dados de contagem, porém não tem como pressuposto que sua média e sua variância possuam os mesmos valores, como na Distribuição de Poisson.

Além do Índice de Fisher para a superdispersão podemos observar que pela análise exploratória das variáveis acidentes de trânsito com vítimas, colisões com vítimas e atropelamentos, apresentada no quadro 7 que as suas respectivas médias são bem inferiores às suas variâncias.

Quadro 7: Estatísticas descritivas (mínimo, máximo, média e variância) das variáveis acidentes de trânsito com vítimas, colisões com vítimas e atropelamentos.

Variáveis	Mínimo	Máximo	Média	Variância
Acidentes de Trânsito com Vítimas	1	2546	224,70	138228,4
Colisões com Vítimas	0	1149	99,67	27435,5
Atropelamentos	0	1665	125,73	48001,7

5.3.1. Acidentes de Trânsito com Vítimas

O processo de modelagem, dos acidentes de trânsito com vítimas, iniciou com a análise da contribuição individual de cada uma das variáveis (pop, area, sanea, menor18, casado, superior, renda, velocmed, via_ap, via_as, via_c, via_e, via_l, lin_bus e intersec) para o modelo. Foram ajustados, então, quinze modelos onde foram observadas as significâncias dos coeficientes dos parâmetros e também a deviance, para se analisar se a variável era significativa para o modelo. Utilizou-se o nível de significância de 5% e assim as variáveis que apresentaram p-valor superior a 5% foram consideradas não significativas. Na análise da deviance foi utilizado a Distribuição de Qui-quadrado com 5% de significância e 1 grau de liberdade ($\chi^2_{0,05;1} = 3,841$). Assim se a contribuição da variável na deviance fosse inferior a 3,841, a variável não deveria ser incluída no modelo.

Após essa primeira análise, foi ajustado um modelo considerando todas as variáveis em estudo e foi também analisada a significância dos coeficientes dos parâmetros e também a deviance para se verificar a contribuição de cada uma das variáveis no modelo.

Através dessas análises verificou-se que as variáveis casado, superior, renda e velocmed não foram significativas na análise individual e também não foram significativas na análise do modelo considerando todas as variáveis.

Foi então construído um modelo inicial onde constavam as variáveis pop, area, sanea, menor18, via_ap, via_as, via_c, via_e, via_l, lin_bus e intersec. Esse modelo identificou que as variáveis via-ap, area, via_e e sanea não contribuíam de forma significativa e, portanto foram excluídas.

Em função dos critérios de não rejeição das variáveis chegou-se ao modelo com as seguintes variáveis explicativas: pop, menor18, via_as e lin_bus. Contudo, as variáveis via_c, via_l e intersec embora não tenham apresentado coeficientes significativos, apresentaram deviance acima de 3,841. Optou-se por analisar um novo modelo onde somente a variável intersec não foi incluída, pois além de não apresentar coeficiente significativo, foi aquela que apresentou a menor contribuição na deviance.

Através da análise do novo modelo verificou-se que a variável via_l não apresentou coeficiente significativo, além de deviance inferior a 3,841. Sendo assim, a mesma foi retirada do modelo.

O modelo que melhor se ajustou foi o que continha as variáveis explicativas via_c, pop, lin_bus, via_as e menor18. Neste modelo todas as variáveis apresentaram coeficientes significativos, além de contribuição na deviance acima de 3,841.

Os possíveis pares interações entre as variáveis via_c, pop, lin_bus, via_as e menor18 foram analisadas formando os seguintes modelos:

- a) via_c*pop+lin_bus+via_as+menor18);
- b) via_c*lin_bus+pop+via_as+menor18;

- c) $\text{via_c} * \text{via_as} + \text{pop} + \text{lin_bus} + \text{menor18}$;
- d) $\text{via_c} * \text{menor18} + \text{pop} + \text{lin_bus} + \text{via_as}$;
- e) $\text{via_c} + \text{pop} * \text{lin_bus} + \text{via_as} + \text{menor18}$;
- f) $\text{via_c} + \text{pop} * \text{via_as} + \text{lin_bus} + \text{menor18}$;
- g) $\text{via_c} + \text{pop} * \text{menor18} + \text{lin_bus} + \text{via_as}$;
- h) $\text{via_c} + \text{pop} + \text{lin_bus} * \text{via_as} + \text{menor18}$;
- i) $\text{via_c} + \text{pop} + \text{lin_bus} * \text{menor18} + \text{via_as}$;
- j) $\text{via_c} + \text{pop} + \text{lin_bus} + \text{via_as} * \text{menor18}$.

Somente nos modelos com interação ($\text{via_c} * \text{pop} + \text{lin_bus} + \text{via_as} + \text{menor18}$) e ($\text{via_c} + \text{pop} * \text{lin_bus} + \text{via_as} + \text{menor18}$) todas as variáveis apresentaram coeficientes significativos e contribuição na deviance acima de 3,841.

Foi analisado um outro modelo conjunto com essas interações ($\text{via_c} * \text{pop}$, $\text{pop} * \text{lin_bus}$, via_as , menor18). Nesse modelo a interação $\text{via_c} * \text{pop}$ apresentou coeficiente não significativo e a interação $\text{pop} * \text{lin_bus}$ apresentou deviance inferior a 3,841.

Foram, então, comparados os seguintes modelos com interação: ($\text{via_c} * \text{pop} + \text{lin_bus} + \text{via_as} + \text{menor18}$) e ($\text{via_c} + \text{pop} * \text{lin_bus} + \text{via_as} + \text{menor18}$).

Para a comparação entre os modelos foram analisadas as significâncias dos parâmetros do modelo, e se os respectivos desvios estavam acima do valor crítico de uma qui-quadrado com 5% de significância e 1 grau de liberdade. Verificou-se também se o desvio residual do modelo (Residual Deviance) estava abaixo de valor crítico de uma qui-quadrado com 5% de significância e com o mesmo número de graus de

liberdade dos resíduos. O critério de informação de Akaike (Akaike Information Criterion) foi utilizado para comparação dos modelos.

A adequação do modelo foi verificada pela análise gráfica, embasada principalmente pelo gráfico Normal de Probabilidade (Normal Probability Plot) e pelo gráfico dos resíduos versus os valores ajustados.

Quando a distribuição do erro é Normal, então o gráfico Normal de Probabilidades se parecerá com uma linha reta onde a maioria dos pontos estará distribuída no meio da reta. Com esse gráfico verifica-se a suposição de Normalidade dos Resíduos.

No gráfico dos resíduos versus os valores ajustados, se o modelo está correto e as suposições estão satisfeitas, os resíduos não devem apresentar qualquer tipo de estrutura, ou seja, eles devem ser não correlacionados com qualquer outra variável, incluindo y . Quando a variabilidade aumenta a medida em que aumenta a magnitude das observações é sinal de que a variância não é constante (heterocedasticidade).

No quadro 8 estão descritos as estimativas dos coeficientes, o erro padrão, a estatística t e o p -valor das variáveis explicativas do modelo que melhor se ajustou aos dados de acidentes de trânsito com vítimas.

Quadro 8: Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo para acidentes de trânsito com vítimas.

Variáveis	Coeficiente	Erro Padrão	t	p-valor	Risco Relativo
Intercepto	4,965e+00	3,399e-01	14,609	< 2e-16	-
via_c	3,225e-03	1,064e-03	3,030	0,00288	1,00323
pop	2,033e-05	4,108e-06	4,949	1,98e-06	1,00002
lin_bus	5,018e-03	9,272e-04	5,413	2,41e-07	1,00503
via_as	6,690e-03	2,169e-03	3,084	0,00243	1,00671
menor18	-4,899e+00	1,162e+00	-4,215	4,30e-05	0,00745
pop:lin_bus	-7,435e-08	1,744e-08	-4,263	3,55e-05	1,00000

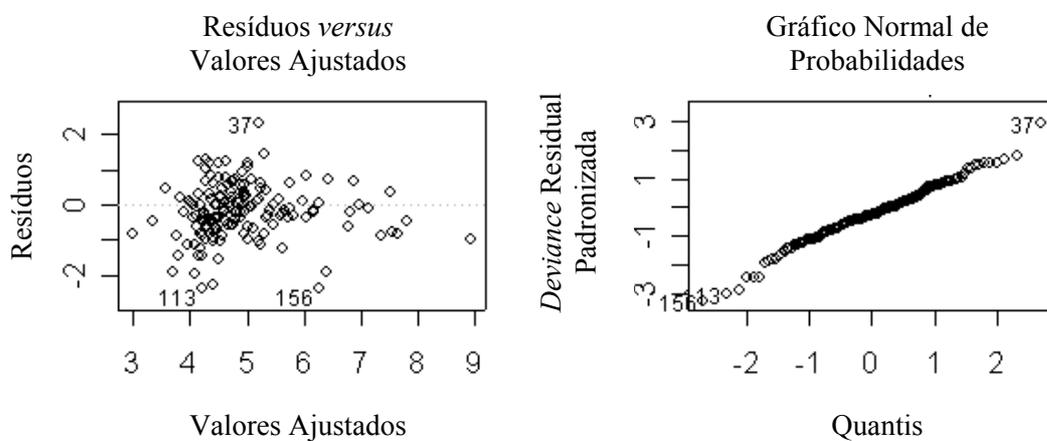


Figura 31: Análise de resíduos para o modelo dos acidentes de trânsito com vítimas com interações via_c*pop+lin_bus+via_as+menor18.

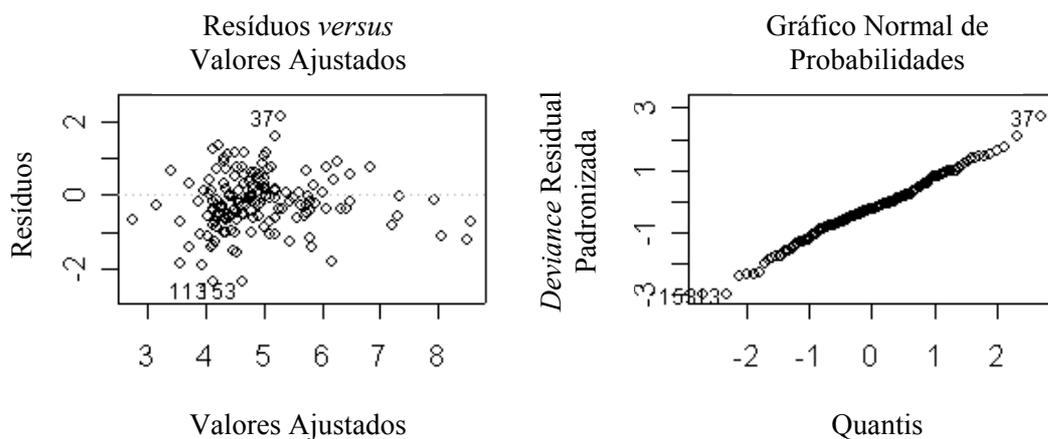


Figura 32: Análise de resíduos para o modelo dos acidentes de trânsito com vítimas com interações via_c+pop*lin_bus+via_as+menor18.

O modelo que melhor se ajustou foi o que continha as seguintes variáveis e interações via_c+pop*lin_bus+via_as+menor18.

$$\hat{y} = \exp(4,965 + 0,003225 \text{ via_c} + 0,00002033 \text{ pop} + 0,005018 \text{ lin_bus} + 0,006690 \text{ via_as} - 4,899 \text{ menor18} - 0,0000007435 \text{ pop} * \text{lin_bus})$$

A escolha desse modelo deve-se ao fato que todas as variáveis explicativas incluídas e a interação foram significantes e também apresentaram desvios acima do valor crítico de uma $\chi^2_{0,05;1} = 3,841$. O desvio residual do modelo está abaixo do valor crítico de uma $\chi^2_{0,05;gl \text{ resíduos}}$, e foi inferior ao do outro modelo estudado. O critério de informação de Akaike (Akaike Information Criterion) foi menor para esse modelo. O número de iterações até a convergência das estimativas dos parâmetros foi maior que o do outro modelo comparado.

5.3.2. Atropelamentos

Na modelagem dos atropelamentos iniciou-se o processo pela análise da contribuição individual de cada uma das variáveis explicativas do estudo. E a análise se deu em função da significância dos parâmetros e também da deviance. Além da contribuição nos modelos individuais foi analisado um modelo contendo as quinze variáveis explicativas.

Através dessas análises verificou-se que as variáveis sanea, velocmed, renda e superior não foram significativas na análise individual e também não foram significativas na análise do modelo considerando todas as variáveis.

Foi construído um modelo inicial onde constavam as variáveis via_c, pop, intersec, lin_bus, via_ap, via_l, via_as, area, via_e, menor18 e casado. A partir desse modelo iniciou-se o processo de modelagem. Esse modelo identificou que as variáveis via_ap, area, via_e e casado não contribuíam de forma significativa e portanto foram excluídas.

Restaram então as variáveis via_c, pop, intersec, lin_bus, via_l, via_as e menor18. Na análise conjunta dessas variáveis verificou-se que via_c, intersec e via_l apresentaram coeficientes não significativos, porém com deviance superior a 3,841. Assim, foi analisado um novo modelo com a exclusão da variável via_l pois além de apresentar coeficiente não significativo, foi a que apresentou menor contribuição na deviance entre as não significativas.

O novo modelo analisado foi composto pelas variáveis via_c, pop, intersec, lin_bus, via_as e menor18. As variáveis via_c e intersec continuaram apresentando coeficientes não significativos, porém a contribuição na deviance também continuou

acima do valor crítico de 3,841. Novamente adotou-se o critério de excluir aquela que apresentou a menor contribuição na deviance. Sendo assim a variável *intersec* foi excluída do modelo.

No modelo composto pelas variáveis *via_c*, *pop*, *lin_bus*, *via_as* e *menor18* continuou sendo observado a não significância da variável *via_c*, mesmo apresentando alta contribuição na deviance e por esse motivo essa variável foi excluída.

As variáveis *pop*, *lin_bus*, *via_as* e *menor18* apresentam coeficientes significativos e contribuição na deviance acima de 3,841.

Foram analisados os pares de interações para o modelo com as variáveis *pop*, *lin_bus*, *via_as* e *menor18*, formando as seguintes interações:

- a) ($pop * lin_bus + via_as + menor18$);
- b) ($pop * via_as + lin_bus + menor18$);
- c) ($pop * menor18 + lin_bus + via_as$);
- d) ($pop + lin_bus * via_as + menor18$);
- e) ($pop + lin_bus * menor18 + via_as$);
- f) ($pop + lin_bus + via_as * menor18$).

Somente os modelos com interações ($pop * lin_bus + via_as + menor18$) e ($pop * via_as + lin_bus + menor18$) apresentaram todos os coeficientes significativos e contribuição na deviance superior a 3,841.

Os três modelos abaixo foram analisados:

- a) $pop + lin_bus + via_as + menor18$;

b) $\text{pop} * \text{lin_bus} + \text{via_as} + \text{menor18}$;

c) $\text{pop} * \text{via_as} + \text{lin_bus} + \text{menor18}$.

O modelo que melhor se ajustou foi $\text{pop} * \text{lin_bus} + \text{via_as} + \text{menor18}$.

No quadro 9 estão descritos as estimativas dos coeficientes, o erro padrão, a estatística t e o p-valor das variáveis explicativas do modelo que melhor se ajustou aos dados de atropelamentos.

Quadro 9: Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo para atropelamentos.

Variáveis	Coefficiente	Erro Padrão	t	p-valor	Risco Relativo
Intercepto	4,312e+00	3,490e-01	12,355	< 2e-16	-
pop	2,759e-05	3,458e-06	7,979	3,41e-13	1,00003
lin_bus	5,675e-03	9,386e-04	6,046	1,12e-08	1,00569
via_as	4,805e-03	2,185e-03	2,199	0,0294	1,00482
menor18	-4,916e+00	1,193e+00	-4,121	6,19e-05	0,00733
pop:lin_bus	-7,547e-08	1,784e-08	-4,231	4,02e-05	1,00000

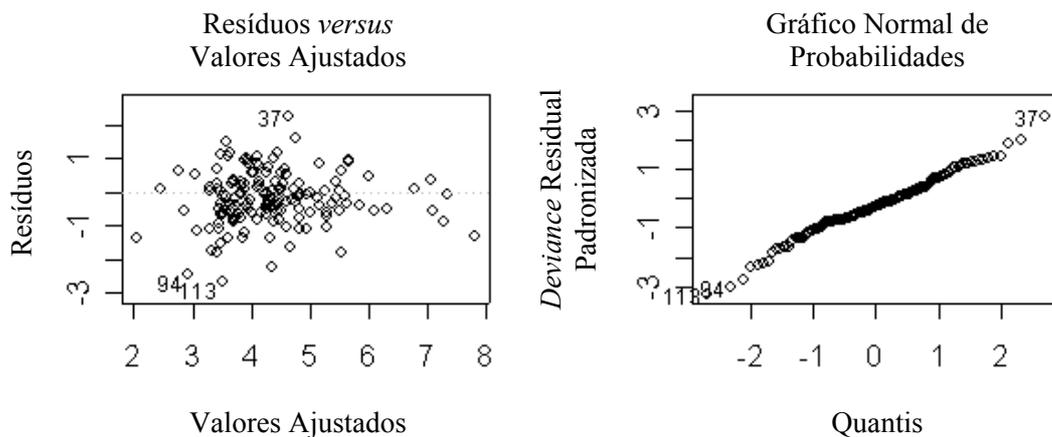


Figura 33: Análise de resíduos para o modelo dos atropelamentos com interações $\text{pop}*\text{lin_bus}+\text{via_as}+\text{menor18}$.

Entre os modelos testados o que demonstrou ser o melhor para descrever o número de atropelamentos, para os bairros do Município do Rio de Janeiro foi:

$$\hat{y} = \exp(4,312 + 0,00002759 \text{ pop} + 0,005675 \text{ lin_bus} + 0,004805 \text{ via_as} + - 4,916 \text{ menor18} - 0,00000007547 \text{ pop} * \text{lin_bus})$$

Neste modelo todas as variáveis explicativas incluídas e a interação apresentaram coeficientes significativos. Além disso, os seus respectivos desvios estavam acima do valor crítico de uma $\chi^2_{0,05;1} = 3,841$. O desvio residual do modelo ficou abaixo do valor crítico de uma $\chi^2_{0,05;gl \text{ resíduos}}$ e o critério de informação de Akaike (AIC) foi o menor entre os modelos testados.

5.3.3. Colisões com Vítimas

Para a modelagem das colisões com vítimas, primeiramente foi analisada a contribuição individual de cada uma das variáveis explicativas do estudo. Essa contribuição foi analisada em função da significância dos parâmetros e também da deviance nos modelos individuais e no modelo conjunto com todas as variáveis explicativas.

Verificou-se que as variáveis *velocmed* e *casado* não foram significativas na análise individual e também não foram significativas na análise do modelo considerando todas as variáveis.

Foi construído um modelo inicial onde constavam as variáveis *via_c*, *pop*, *intersec*, *via_l*, *lin_bus*, *via_e*, *area*, *via_ap*, *via_as*, *menor18*, *renda*, *sanea* e *superior*. Esse modelo identificou que as variáveis *via_e*, *area*, *via_ap*, *renda*, *sanea* e *superior* não contribuíam de forma significativa e portanto, essas variáveis foram excluídas.

Um modelo com as variáveis *via_c*, *pop*, *intersec*, *via_l*, *lin_bus*, *via_as* e *menor18* foi analisado e verificou-se que todas apresentaram deviance superior a 3,841 e que as variáveis *via_c*, *pop*, *intersec* e *via_l* não apresentaram coeficientes dos parâmetros significativos. Foi então construído um novo modelo sem a variável *intersec*, pois entre aquelas que apresentaram coeficiente não significativo, está foi a que apresentou menor contribuição na deviance.

No novo modelo foram analisadas as variáveis *via_c*, *pop*, *via_l*, *lin_bus*, *via_as* e *menor18*. Nesta análise observou-se que as variáveis *pop* e *via_l* embora não apresentassem coeficientes significativos, apresentaram deviance superior a 3,841, e

entre elas a variável pop foi a que menor contribuiu para a deviance sendo então excluída.

Foi analisado o modelo com as variáveis via_c, via_l, lin_bus, via_as e menor18. Nesta análise a variável via_l não apresentou coeficiente significativo e a sua deviance foi inferior a 3,841.

O modelo com as variáveis via_c, lin_bus, via_as e menor18 foi o que apresentou coeficientes significativos além de contribuição na deviance acima de 3,841.

Foram então, analisados todos os pares de interações para o modelo com as variáveis via_c, lin_bus, via_as e menor18, formando os seguintes modelos com interações:

- a) $\text{via_c} * \text{lin_bus} + \text{via_as} + \text{menor18}$;
- b) $\text{via_c} * \text{via_as} + \text{lin_bus} + \text{menor18}$;
- c) $\text{via_c} * \text{menor18} + \text{lin_bus} + \text{via_as}$;
- d) $\text{via_c} + \text{lin_bus} * \text{via_as} + \text{menor18}$;
- e) $\text{via_c} + \text{lin_bus} * \text{menor18} + \text{via_as}$;
- f) $\text{via_c} + \text{lin_bus} + \text{via_as} * \text{menor18}$.

Nos modelos com interação ($\text{via_c} * \text{lin_bus} + \text{via_as} + \text{menor18}$), ($\text{via_c} * \text{via_as} + \text{lin_bus} + \text{menor18}$), ($\text{via_c} + \text{lin_bus} * \text{via_as} + \text{menor18}$) e ($\text{via_c} + \text{lin_bus} + \text{via_as} * \text{menor18}$) todas as variáveis apresentaram coeficientes significativos e contribuição na deviance acima de 3,841.

Como os modelos com todas as combinações possíveis de duas variáveis para via_c, lin_bus e via_as apresentaram coeficientes significativos e contribuição na

deviance acima de 3,841 foi analisado modelo com interação tripla ($\text{via_c}*\text{lin_bus}*\text{via_as}+\text{menor18}$) e também o modelo com uma interação tripla e outra dupla ($\text{via_c}*\text{lin_bus}*\text{via_as}+\text{via_as}*\text{menor18}$). Esses modelos não apresentaram coeficientes significativos além da contribuição na deviance ser inferior ao valor de uma Qui-Quadrado com 1 grau de liberdade.

Foram então, analisados os cinco modelos abaixo relacionados:

- a) $\text{via_c}+\text{lin_bus}+\text{via_as}+\text{menor18}$;
- b) $\text{via_c}*\text{lin_bus}+\text{via_as}+\text{menor18}$;
- c) $\text{via_c}*\text{via_as}+\text{lin_bus}+\text{menor18}$;
- d) $\text{via_c}+\text{lin_bus}*\text{via_as}+\text{menor18}$;
- e) $\text{via_c}+\text{lin_bus}+\text{via_as}*\text{menor18}$.

O modelo que melhor se ajustou foi $\text{via_c}*\text{via_as}+\text{lin_bus}+\text{menor18}$.

No quadro 10 estão descritos as estimativas dos coeficientes, o erro padrão, a estatística t e o p-valor das variáveis explicativas do modelo que melhor se ajustou aos dados de colisões com vítimas.

Quadro 10: Estimativas dos coeficientes, erro padrão, estatística t, p-valor e risco relativo das variáveis explicativas do modelo para colisões com vítimas.

Variáveis	Coefficiente	Erro Padrão	t	p-valor	Risco Relativo
Intercepto	4,461e+00	3,424e-01	13,028	< 2e-16	-
via_c	6,763e-03	6,712e-04	10,075	< 2e-16	1,00679
via_as	1,828e-02	3,095e-03	5,907	2,22e-08	1,01845
lin_bus	2,733e-03	7,179e-04	3,807	0,000204	1,00274
menor18	-5,004e+00	1,189e+00	-4,208	4,40e-05	0,00671
via_c:via_as	-6,627e-05	1,593e-05	-4,159	5,33e-05	0,99993

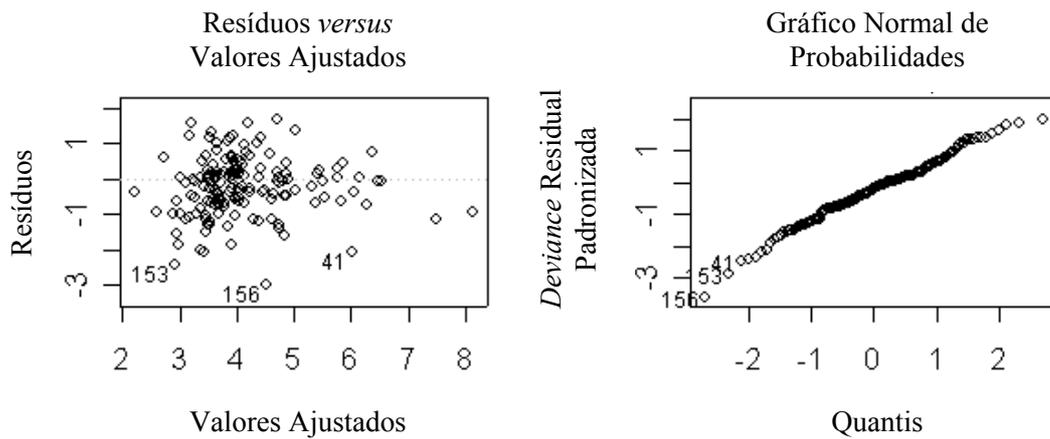


Figura 34: Análise de resíduos para o modelo das colisões com vítimas com interações via_c*via_as+lin_bus+menor18.

Entre os modelos testados o que demonstrou ser o melhor para descrever o número de colisões com vítimas, para os bairros do Município do Rio de Janeiro foi:

$$\hat{y} = \exp(4,461 + 0,006763 \text{ via_c} + 0,01828 \text{ via_as} + 0,002733 \text{ lin_bus} + \\ - 5,004 \text{ menor18} - 0,00006627 \text{ via_c} * \text{via_as})$$

A escolha desse modelo deve-se ao fato que todas as variáveis explicativas incluídas e a interação apresentaram coeficientes significativos além de serem significantes devido ao seus respectivos desvios estarem acima do valor crítico de uma $\chi^2_{0,05;1} = 3,841$. Além disso, o desvio residual do modelo está abaixo do valor crítico de uma $\chi^2_{0,05;gl \text{ resíduos}}$. O critério de informação de Akaike foi o menor entre os modelos testados.

6. DISCUSSÃO

Estudos sobre acidentes de trânsito ainda são poucos em nosso país. Esse grupo de causas específicas que junto com os homicídios, suicídios, quedas, queimaduras, afogamentos entre outras compõe o capítulo 20 (Causas Externas de Morbidade e Mortalidade), da 10ª revisão da Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde (OMS, 1996), vem sendo pouco estudado uma vez que na maioria dos grupos etários não é a principal causa de mortes no conjunto das causas externas, sendo superado pelos homicídios.

Os poucos estudos nacionais que abordam este tema são, na maioria das vezes, descritivos, não conseguindo analisar de forma mais profunda o impacto que esse tipo de causa pode desencadear na população e quais são os principais fatores desencadeantes desse quadro para que se possa pensar em políticas que visem a diminuição no número de ocorrências e nas taxas de morbidade e mortalidade por acidentes de trânsito no Brasil.

Um fator preocupante é o problema da subnotificação, apontada nos estudos de LADEIRA e GUIMARÃES (1998), ANDRADE e MELLO-JORGE (2001), BARROS et al. (2003), que pode ocorrer, em geral, por falta do registro ou até mesmo pela sua inclusão em outro grupo de causas.

Estudos como os de ANDRADE e MELLO-JORGE (2001), BARROS et al. (2003) e BASTOS et al. (2005) mostraram que a maior concentração de casos de acidentes de trânsito ocorre nos finais de semana e este mesmo comportamento pode ser observado nesse estudo através da análise descritiva básica (figura 4).

Neste trabalho buscou-se inicialmente analisar a estrutura e a distribuição espacial do número de acidentes de trânsito envolvendo vítimas, ocorridos no período de 01/01/1998 a 31/12/2000, nos bairros do Município do Rio de Janeiro, registrados pela Polícia Civil.

Para essa análise procurou-se, primeiramente, identificar as principais áreas de incidência e as suas relações com os vizinhos através da construção de mapas cloropléticos e do cálculo de indicadores espaciais da dependência espacial entre áreas vizinhas.

A expectativa era a criação de estimadores espaciais que incorporassem a dependência espacial entre as áreas vizinhas, que são de grande auxílio na detecção de áreas mais problemáticas, contribuindo assim para o planejamento de medidas que possibilitem o controle e a diminuição na ocorrência desses eventos, e também para a caracterização epidemiológica deste agravo de saúde.

Contudo, os dados demonstraram não haver dependência espacial entre as áreas de estudo, tanto para o número total de acidentes de trânsito com vítimas, assim como para o total de atropelamentos e o total de colisões com vítimas, o que impossibilitou a continuação do trabalho através dessa abordagem, não podendo ser agregada aos modelos em estudo a componente espacial.

A falta de dependência espacial pode estar relacionada ao fato dos agrupamentos espaciais em áreas maiores afetar a habilidade em se detectar agregados de doenças, ou investigar relações suspeitas entre quadros ambientais e eventos de doença (ARMSTRONG et al., 1999, BOULOS, 2004). Contudo como não se conseguiu um melhor detalhamento da informação de modo que pudessem ser utilizados como

unidades geográficas os setores censitários do município do Rio de Janeiro, ou até mesmo a informação georeferenciada por ponto, optou-se por trabalhar com os registros agregados em bairros. PETCH e HENSON (2000) em seu estudo com crianças pedestres e ciclistas, embora pudessem utilizar como unidades geográficas os distritos da Cidade de Salford, no Reino Unido, optaram por agrupar os distritos formando áreas maiores uma vez que o número de casos observados não foi suficientemente grande para abranger os 485 distritos da cidade. Esses autores destacam que a agregação baseada somente na localização dos distritos não foi adequada em função das muitas variações existentes entre eles em termos das características sócio-econômicas, de tráfego e ambientais.

O agrupamento adotado nesse estudo pode não ter sido o que melhor se enquadrava, porém foi único possível devido às limitações do banco de dados já descritas anteriormente. Assim, como o Município do Rio de Janeiro possui vários contrastes diferentes em um mesmo bairro, o ideal seria trabalhar com essa informação em um nível menor de agregação buscando identificar as especificidades de cada uma dessas áreas e as suas influências com a ocorrência dos acidentes de trânsito com vítimas.

Buscando identificar as principais variáveis facilitadoras da ocorrência de acidentes de trânsito foram construídos modelos de regressão Binomial Negativa para o número de casos de acidentes de trânsito com vítimas, o número de atropelamentos e o número de colisões com vítimas. Verificou-se que distribuição de Poisson não seria a melhor opção a ser empregada para a construção dos modelos, pois foi observada superdispersão nos dados. A utilização da distribuição Binomial Negativa ocorreu em função do fenômeno da superdispersão ser muito comum nos dados de acidentes de

trânsito, como destacam MIAOU e LUM (1993) e AMOROS et al. (2003), violando um pressuposto para a utilização do modelo de regressão de Poisson e acarretando em uma super ou subestimação dos erros padrões, que podem propiciar uma avaliação errônea da significância dos parâmetros de regressão.

O modelo de regressão Binomial Negativa mostrou que o número de vias coletoras no bairro, a população, o número de linhas de ônibus que trafegam no bairro, o número de vias arterial secundária no bairro e a proporção de pessoas com idade inferior a 18 anos dos bairros do Município do Rio de Janeiro são variáveis importantes para representar o número total de casos de acidentes de trânsito com vítimas. As variáveis número de vias coletoras, população, número de linhas de ônibus e número de vias arterial secundária apresentaram uma relação positiva com o evento, enquanto a variável proporção de menores de 18 anos e a interação população versus o número de linhas de ônibus têm uma relação negativa, com a ocorrência desse evento. Embora entre os modelos testados, esse demonstrou ser o mais adequado, algumas considerações devem ser feitas a respeito das suas estimativas. Esse modelo superestimou o número de acidentes de trânsito com vítimas nos bairros onde foram observados os menores números de casos. Dos quinze bairros com o menor número de ocorrências em doze deles houve superestimação que variou na ordem de 276,9% a 6100,0%, respectivamente em Pitangueiras, com 13 casos observados e 49 casos estimados e na Praia da Bandeira, com 1 caso observado e 62 casos estimados. As ocorrências dos acidentes de trânsito com vítimas, nos bairros de Santa Cruz e Tijuca também foram superestimadas pelo modelo, embora esses bairros estejam entre os com maiores números de casos.

Ao analisar o risco relativo das variáveis significativas no modelo para acidentes de trânsito com vítimas no município do Rio de Janeiro (quadro 8) observa-se que o número de vias arterial secundária é a variável que apresenta maior risco relativo ($RR = 1,00671$) enquanto a variável proporção de menores de 18 anos é a que apresenta o menor risco relativo ($RR = 0,00745$), aparentando então ser um fator de proteção para a ocorrência de acidentes de trânsito com vítimas, ou seja, os bairros que possuem as maiores proporções de pessoas com idade inferior a 18 anos, em geral, são aqueles que possuem os menores números de casos de acidentes de trânsito com vítimas no município do Rio de Janeiro, pois o risco relativo maior que um significa que um aumento no valor da variável implica em um aumento no número de casos e risco relativo menor que um significa que um aumento no valor da variável implica em uma diminuição no número de casos (OLMSTEAD, 2001, CHIN e QUDDUS, 2003). A única exceção foi verificada no bairro de Santa Cruz, que embora tenha apresentado uma das maiores proporções de pessoas com idade inferior a 18 anos, apresentou também um grande número de casos de acidentes de trânsito com vítimas. É importante destacar que esse bairro possui a terceira maior população do município e é o segundo maior em número de vias coletoras. A variável população e a interação população versus o número de linhas de ônibus embora tenham sido consideradas significativas pelo modelo, apresentaram risco relativo de aproximadamente um, não sendo possível caracterizá-las como “fator de risco” ou “fator de proteção” para a ocorrência do evento.

Os atropelamentos ocorridos nos bairros do município do Rio de Janeiro foram descritos, pelo modelo de regressão Binomial Negativa, com as variáveis: população; número de linhas de ônibus que trafegam no bairro; número de vias arterial secundária; proporção de pessoas com idade inferior a 18 anos e a interação população versus

número de linhas de ônibus. As variáveis população, número de linhas de ônibus que trafegam no bairro e número de vias arterial secundária têm uma relação positiva, e a variável proporção de pessoas com idade inferior a 18 anos e a interação população versus número de linhas de ônibus têm uma relação negativa, com a ocorrência desse evento. O modelo para atropelamentos também superestimou essas ocorrências nos bairros com o menor número de casos. Entre os dezessete bairros com o menor número de ocorrências em dezesseis deles houve superestimação variando de 210,0% no Vidigal, que apresentou 10 ocorrências e o modelo estimou 31, a 3850,0% no Complexo do Alemão, com 2 casos observados e 79 casos estimados. Entre os bairros com maior número de ocorrências, Tijuca e Copacabana foram, em termos proporcionais, os que apresentaram maior superestimação pelo modelo. Esses bairros estão entre os que possuem maior população e maior número de vias arterial secundárias.

No modelo para o número de atropelamentos no município do Rio de Janeiro o maior risco relativo (quadro 9) foi observado para a variável o número de linhas de ônibus que trafegam no bairro ($RR = 1,00569$) e como menor risco relativo a variável proporção de menores de 18 anos ($RR = 0,00733$). Como no modelo para o número de acidentes de trânsito com vítimas no município do Rio de Janeiro, a variável população e a interação população versus o número de linhas de ônibus apresentaram risco relativo de aproximadamente um.

O número de vias coletoras no bairro, o número de vias arterial secundárias no bairro, o número de linhas de ônibus que trafegam no bairro, e a proporção de pessoas com idade inferior a 18 anos dos bairros do Município do Rio de Janeiro são variáveis importantes para descrever as colisões com vítimas, ocorridos nessas áreas. As variáveis

número de vias coletoras, número de vias arterial secundária e número de linhas de ônibus apresentaram uma relação positiva com o evento, e a variável proporção de menores de 18 anos e a interação número de vias coletoras versus número de vias arterial secundárias têm uma relação negativa, com a ocorrência desse evento. O modelo para colisões com vítimas também superestimou as ocorrências, principalmente nos bairros onde o número de casos observados foi menor. Nos bairros de Santa Cruz e Campo Grande, que estão entre os que possuem os maiores números de colisões com vítimas, o valor estimado também foi bem superior ao observado e esses bairros são os que possuem o maior número de vias coletoras.

A variável número de vias arterial secundárias foi a que apresentou, conforme o quadro 10, o maior risco relativo no modelo para o número de colisões com vítimas no município do Rio de Janeiro. Como aconteceu nos modelos anteriores a variável proporção de menores de 18 anos foi a que apresentou o menor risco relativo (RR = 0,00671). A interação entre as variáveis número de vias coletoras e número de vias arterial secundárias apresentou risco relativo de aproximadamente um.

Na análise dos modelos selecionados verificou-se que a variável população não foi significativa no modelo para colisões com vítimas e que a variável quantidade total de vias coletoras inseridas em cada da área (bairro) não apresentou coeficiente significativo no modelo para atropelamentos. No modelo para acidentes de trânsito com vítimas essas duas variáveis foram significativas. Isso corrobora a afirmação de SHANKAR et al. (1995) que sugerem que os modelos separados segundo tipo de acidente são melhores em captar as especificidades que modelos para acidentes de trânsito em geral.

Esse comportamento parece ser mais razoável, pois se deve esperar que em um modelo que procure representar o número de casos de atropelamentos em determinada localidade se tenha como variável explicativa a população exposta. Da mesma forma, em um modelo que procure representar o número de colisões com vítimas parece ser mais razoável ter como variável explicativa a quantidade total de vias coletoras inseridas em cada área (bairro), uma vez que são vias destinadas a coletar o tráfego das vias locais e distribuí-lo nas vias arteriais e expressas, a servir de rota de transporte coletivo e a atender na mesma proporção o tráfego de passagem e local com razoável padrão de fluidez.

Embora LIMA e VASCONCELOS (1999) tenham verificado um aumento nas taxas de mortalidade por acidentes de trânsito nos períodos de aumento da atividade econômica, especificamente, em 1986 (Plano Cruzado) e 1994 (Plano Real), a análise dos modelos para acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas revelam que o fator econômico não apresentou relação com a ocorrência do evento. A renda mediana da área não apresentou coeficiente significativo em nenhum dos modelos selecionados. Além disso, outras variáveis que poderiam estar relacionadas ao padrão econômico como nível superior de escolaridade e saneamento básico também não apresentaram significância estatística.

O fato dos modelos aqui apresentados para a estimação do número de acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas superestimarem o número de ocorrências nos bairros de menor incidência pode estar diretamente relacionado ao problema da subnotificação.

No Complexo do Alemão e no Complexo da Maré o número de ocorrências de acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas foram superestimados.

O Complexo do Alemão é um bairro relativamente novo, tendo sido criado em 1993, que possui como vizinhos os bairros de Bonsucesso, Ramos, Olaria, Higienópolis, Inhaúma e Engenho da Rainha. Para os bairros de Bonsucesso e Inhaúma foi observada uma subestimação do número de ocorrências de acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas. Como os dados utilizados nesse trabalho referem-se aos anos de 1998 a 2000 e em função do ano de criação do bairro, os casos ocorridos no Complexo do Alemão podem ter sido incluídos em seus vizinhos, principalmente em Bonsucesso e Inhaúma.

O comportamento descrito para o Complexo do Alemão também foi observado no Complexo da Maré cuja criação ocorreu em 1994. O Complexo da Maré possui como vizinhos os bairros Manguinhos, Bonsucesso, Ramos, Olaria, Penha, Cidade Universitária e Caju. O número de ocorrências de acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas foi subestimado pelos modelos em Manguinhos, Bonsucesso e Penha. Logo os casos ocorridos no Complexo da Maré podem ter sido incluídos em seus vizinhos, principalmente nos bairros onde houve subestimação.

Além da subnotificação, a agregação em bairros também pode estar influenciando a superestimação nas áreas de menor ocorrência, uma vez que as suas especificidades são mais difíceis de serem detectadas quando a agregação ocorre em um nível maior, sendo mais aconselhável trabalhar com espaços menores, o que não foi possível nesse trabalho.

A inclusão de novas variáveis explicativas também poderia diminuir a superestimação observada. O fluxo viário, a extensão e a largura das vias, a quantidade de semáforos, a quantidade escolas, hospitais, mercados entre outros equipamentos sociais, poderiam tornar os resultados apresentados por esses modelos mais próximos da realidade observada.

7. CONCLUSÕES

As análises descritas ao longo do trabalho apontam para a necessidade de implantação de uma estratégia de fiscalização, objetivando a redução no número de ocorrências, que deva ser adotada, principalmente, fora do horário comercial, pois a maioria dos casos aqui estudados ocorreu nos finais de semana.

A utilização do modelo de regressão Binomial Negativa foi proposta para estabelecer as relações existentes entre o número de acidentes de trânsito com vítimas, o número de atropelamentos e o número de colisões com vítimas ocorridos nos bairros do município do Rio de Janeiro e as características socioeconômicas, ambientais e demográficas dessas localidades. Os testes indicaram que o modelo de regressão de Poisson não seria o melhor a ser empregado no estudo, pois foi observada superdispersão na variável resposta.

Deve-se destacar que embora os modelos ajustados para acidentes de trânsito com vítimas, atropelamentos e colisões com vítimas, aqui apresentados, fossem os melhores entre os testados, houve uma superestimação do número de ocorrências, principalmente, nos bairros com menor incidência desses eventos. Essa superestimação pode estar diretamente relacionada a uma subnotificação do número de ocorrências nesses bairros que pode ter sido ocasionada em função da alocação dos registros nos bairros vizinhos, como descrito anteriormente para o Complexo do Alemão e o Complexo da Maré, ou também em função de ter sido utilizada a agregação em bairro, em função da qualidade do registro. Além disso, a inclusão de novas co-variáveis, como por exemplo, o fluxo viário, a extensão e a largura das vias, a quantidade de semáforos,

a quantidade escolas, hospitais, mercados entre outros equipamentos sociais, já descritas nos estudos de ABDEL-ATY e RADWAN (1999), LASCALA (2000) e GREIBE (2003) pode aumentar o poder de explicação dos modelos e diminuir os casos de subestimação e superestimação das ocorrências.

A identificação das principais variáveis relacionadas aos acidentes de trânsito com vítimas atende aos objetivos propostos nesse trabalho e representa um importante passo no estudo desse grupo de causas específicas no conjunto das causas externas, pois poucos trabalhos abordam, de forma mais analítica, esse tema no Brasil.

O fato da variável população não ter apresentado coeficiente significativo no modelo proposto para as colisões com vítimas e a variável quantidade total de vias coletoras inseridas em cada área (bairro) não ser significativa no modelo para atropelamentos, pode ser analisado em função da exposição ao risco, uma vez que quanto maior o número de vias coletoras espera-se uma maior concentração de tráfego, aumentado assim o risco de colisões. De forma análoga, quanto maior o número de pessoas em circulação espera-se que seja maior o risco de atropelamentos. O coeficiente de associação linear entre as variáveis número de colisões com vítimas e número de vias coletoras foi de 77,3% e entre as variáveis número de atropelamentos e população foi de 68,1%.

Os resultados desse estudo podem ser empregados na identificação de áreas de maior ocorrência de acidentes, a partir de características ambientais, socioeconômicas e demográficas e na elaboração de políticas que permitam focar os grupos mais suscetíveis a esses eventos buscando a diminuição no número de acidentes envolvendo vítimas.

A notificação mais precisa do local de ocorrência desses eventos pode viabilizar estudos onde se utilizem informações georeferenciadas e propiciem focar estratégias para o controle. Para isso, é importante que órgãos como a Polícia Civil, Polícia Militar, Corpo de Bombeiros e Secretarias Estaduais e Municipais de Saúde e de Transportes se mobilizem visando à construção de um cadastro mais abrangente e com maior rigor no seu preenchimento.

8. REFERÊNCIAS BIBLIOGRÁFICAS

- ABDEL-ATY, M. A., RADWAN, A. E., 2000, “Modeling Traffic Accident Occurrence and Involvement”, *Accident Analysis and Prevention*, v. 32, pp. 633-642.
- ADURA, F. E., 2002, *Medicina de Tráfego: 101 Perguntas e Respostas*. 1 ed. São Paulo, ABRAMET.
- AMOROS, E., MARTIN, J. L., LAUMON, 2003, “Comparation of Road Crashes Incidence and Severity between some French Counties”, *Accident Analysis and Prevention*, v. 35, pp. 537-547.
- ANDRADE, S. M., MELLO-JORGE, M. H. P., 2001, “Acidentes de Transporte Terrestre em Cidade da Região Sul do Brasil: Avaliação da Cobertura e Qualidade dos Dados”, *Cadernos de Saúde Pública*, v. 17, n. 6, pp. 1449-1456.
- ARMSTRONG, M. P., RUSHTON, G., ZIMMERMAN, D. L., 1999, “Geographically Masking Health Data to Preserve Confidentiality”, *Statistics in Medicine*, v. 18, n. 5, 497-525.
- ASSIS, S. G., SOUZA, E. R., 1995, “Morbidade por Violência em Crianças e Adolescentes do Município do Rio de Janeiro”, *Jornal de Pediatria*, v. 71, n. 6, pp. 303-312.
- BAILEY, T. C., GATRELL, A. C., 1995, *Interactive Spatial Data Analysis*. 1 ed. Essex, Longman Scientific & Technical.

- BARROS, A. J. D., AMARAL, R. L., OLIVEIRA, M. S. B., *et al.*, 2003, “Acidentes de Trânsito com Vítimas: Sub-registro, Caracterização e Letalidade”, *Cadernos de Saúde Pública*, v. 19, n. 4, pp. 979-986.
- BARROS, M. D. A., XIMENES, R., LIMA, M. L. C., 2001, “Mortalidade por Causas Externas em Crianças e Adolescentes: Tendências de 1979 a 1995”, *Revista de Saúde Pública / Journal of Public Health*, v. 35, n. 2, pp. 142-149.
- BASTOS, Y. G. L., ANDRADE, S. M., CORDONI JÚNIOR, L., 1999, “Acidentes de Trânsito e o Novo Código de Trânsito Brasileiro em Cidade da Região Sul do Brasil”, *Informe Epidemiológico do SUS*, v. 8, n. 2, pp. 37-45.
- BASTOS, Y. G. L., ANDRADE, S. M., SOARES, D. A., 2005, “Características dos Acidentes de Trânsito e das Vítimas Atendidas em Serviço Pré-hospitalar em Cidade do Sul do Brasil, 1997/2000”, *Cadernos de Saúde Pública*, v. 21, n. 3, pp. 815-822.
- BERHANU, G., 2004, “Models Relating Traffic Safety with Road Environment and Traffic Flows on Arterial Roads in Addis Ababa”, *Accident Analysis and Prevention*, v. 36, pp. 697-704.
- BOULOS, M. N. K., 2004, “Towards Evidence-based, GIS-driven National Spatial Health Information Infrastructure and Surveillance Services in the United Kingdom”, *International Journal of Health Geographics*, vol. 3, pp. 1-50.
Disponível em: < <http://www.ij-healthgeographics.com/content/3/1/1> >. Acesso em: 24.jan.2006.

BRASIL. Lei n.º 9.503, de 23 de setembro de 1997. Institui o código de trânsito brasileiro. Diário Oficial da República Federativa do Brasil, Brasília, DF, 24 de set. 1997. Disponível em: < <http://www.denatran.gov.br/Legislacao.htm> >. Acesso em: 22.set.2003.

CASTILLO, J., PÉREZ-CASANY, M., 2005, “Overdispersed and Underdispersed Poisson Generalizations”, *Journal of Statistical Planning and Inference*, v. 134, pp. 486-500.

CENEPI/Fundação Nacional de Saúde, 1997, *Mortalidade Brasil 1994*.

CHANG, L.-Y., 2005, “analysis of Freeway Accident Frequencies: Negative Binomial Regression Versus Artificial Neural Network”, *Safety Science*, v. 43, pp. 541-557.

CHIN, H. C., QUDDUS, M. A., 2003, “Applying the Random Effect Negative Binomial to Examine Traffic Accident Occurrence at Signalized Intersections”, *Accident Analysis and Prevention*, v. 35, pp. 253-259.

CORDEIRO, G. M., 1986, *Modelos Lineares Generalizados*. 1 ed. Campinas, VII Simpósio Nacional de Probabilidade e Estatística.

CRESSIE, N. A. C., 1991, *Statistical for Spatial Data*. 1 ed. New York, John Wiley & Sons.

DESLANDES, S. F., SILVA, C. M. F. P., 2000, “Análise da Morbidade Hospitalar por Acidentes de Trânsito em Hospitais Públicos do Rio de Janeiro, RJ, Brasil”, *Revista de Saúde Pública / Journal of Public Health*, v. 34, n. 4, pp. 367-372.

DOBSON, A. J., 1983, *An Introduction to Statistical Modelling*. 1 ed. New York, Chapman and Hall.

- DRAPER, N. R., SMITH, H., 1981, *Applied Regression Analysis*. 2 ed. New York, John Wiley & Sons.
- DUARTE, E. C., SCHNEIDER, M. C., PAES-SOUSA, R., *et al.*, 2002, *Epidemiologia das Desigualdades em Saúde no Brasil: Um Estudo Exploratório*. 1 ed. Brasília, Organização Pan-Americana de Saúde.
- FRIDSTROM, L., IFVER, J., INGEBRIGTSEN, S., *et al.*, 1995, “Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts”, *Accident Analysis and Prevention*, v. 27, n. 1, pp. 1-20.
- GREIBE, P., 2003, “Accident Prediction Models for Urban Roads”, *Accident Analysis and Prevention*, v. 35, pp. 273-285.
- GRISCI, C. L. J., 1991, “Relação entre Acidente de Trânsito e as Variáveis de Agressividade, Atuação e Culpa”, *Psico*, v. 21, n. 10, pp. 103-117.
- Guia de ruas Páginas Amarelas, 1992, *Guia de Ruas 92 – Grande Rio*. 1 ed. Rio de Janeiro, Editora Expressão e Cultura, Editora Páginas Amarelas.
- Guia Quatro Rodas, 2002, *Guia de Ruas Rio de Janeiro 2002*. 1 ed. Rio de Janeiro, Editora Abril.
- GÜRTLER, N., HENZE, N., 2000, “Recent and Classical Goodness-of-fit Tests for the Poisson Distribution”, *Journal os Statistical Planning and Inference*, v. 90, pp. 207-225.

- HIJAR-MEDINA, M. C., CARRILLO-ORDAZ, C. E., FLORES-ALDANA, M. E., *et al.*, 1999, “Factores de Riesgo de Lesión por Accidentes de Tráfico y el Impacto de una Intervención sobre la Carretera”, *Revista de Saúde Pública / Journal of Public Health*, v. 33, n. 4, pp. 504-512.
- HOFFMANN, R., VIEIRA, S., 1998, *Análise de Regressão. Uma Introdução à Econometria*. 3 ed. São Paulo, Editora HUCITEC.
- HU, P. S., TRUMBLE, D. A., FOLEY, D. J., 1998, “Crash Risks of Older Drivers: A Panel Data Analysis”, *Accident Analysis and Prevention*, v. 30, n. 5, pp. 569-581.
- JEGEDE, F. J., 1988, “Spatio-Temporal Analysis of Road Traffic Accidents in Oyo State, Nigeria”, *Accident Analysis and Prevention*, v. 20, n. 3, pp. 227-243.
- KMET, L., BRASHER, P., MACARTHUR, C., 2003, “A Small Area Study of Motor Vehicle Crash Fatalities in Alberta, Canada”, *Accident Analysis and Prevention*, v. 35, pp. 177-182.
- KOTZ, S., JOHNSON, N. L., 1983, *Encyclopedia of Statistical Sciences*. Vol. 3, 1 ed. New York, John Wiley & Sons.
- KRZANOWSKI, W. J., 1998, *An Introduction to Statistical Modelling*. 1 ed. London, Edward Arnold.
- LADEIRA, R. M., GUIMARÃES, M. D. C., 1998, “Análise da Concordância da Codificação de Causa Básica de Óbito por Acidentes de Trânsito”, *Revista de Saúde Pública / Journal of Public Health*, v. 32, n. 2, pp. 502-507.

- LASCALA, E. A., GERBER, D., GRUENEWALD, P. J., 2000, “Demographic and Environmental Correlates of Pedestrian Injury Collisions: A Spatial Analysis”, *Accident Analysis and Prevention*, v. 32, pp. 651-658.
- LAIT, L. C. F., 1992, “Estimativa da Mortalidade por Homicídios no Estado do Rio de Janeiro de 1979 a 1987 e 1989”. *II Congresso Brasileiro de Epidemiologia*, Belo Horizonte, MG, Brasil, 13-17 julho.
- LEITÃO, F. B. P., 1997, “IV – Relato de Experiência Quanto à Prevenção de Acidentes de Trânsito: Um Modelo Usado pela Organização Panamericana de Saúde”, *Revista de Saúde Pública / Journal of Public Health*, v. 31 (suplemento 4), pp. 47-50.
- LIMA, D. D., 1996, “Óbitos por Acidentes de Trânsito no Distrito Federal, 1980 a 1989”, *Revista de Saúde do Distrito Federal*, v. 7, n. 1, pp. 25-28.
- LIMA, D. D., VASCONCELOS, A. M. N., 1999, “A Mortalidade por Acidentes de Trânsito no Brasil”, *Revista da ABRAMET*, n. 30, pp. 44-51.
- MARIN, L., QUEIROZ, M. S., 2000, “A Atualidade dos Acidentes de Trânsito na Era da Velocidade: Uma Visão Geral”, *Cadernos de Saúde Pública*, v. 16, n. 1., pp. 7-21.
- MELLO JORGE, M. H. P., 1990, “Situação Atual das Estatísticas Oficiais Relativas à Mortalidade por Causas Externas”, *Revista de Saúde Pública / Journal of Public Health*, v. 24, n. 3, pp. 217-223.
- MELLO JORGE, M. H. P., GOTLIEB, S.L.D., LAURENTI, R., 2001, *A Saúde no Brasil: Análise do Período de 1996 a 1999*. 1 ed. Brasília, Organização Pan-Americana de Saúde.

- MELLO JORGE, M. H. P., LATORRE, M. R. D. O., 1994, “Acidentes de Trânsito no Brasil: Dados e Tendências”, *Cadernos de Saúde Pública*, v. 10 (suplemento 1), pp. 19-44.
- MIAOU, S., 1994, “The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions”, *Accident Analysis and Prevention*, v. 26, n. 4, pp. 471-482.
- MIAOU, S., LUM, H., 1993, “Modeling Vehicle Accidents and Highway Geometric Design Relationships”, *Accident Analysis and Prevention*, v. 25, n. 6, pp. 689-709.
- MILLER, T. R., 1993, “Costs and Functional Consequences of U.S. Roadway Crashes”, *Accident Analysis and Prevention*, v. 25, n. 5, pp. 593-607.
- MINAYO, M. C. S., 1994, “A Violência Social sob a Perspectiva da Saúde Pública”, *Cadernos de Saúde Pública*, v. 10 (suplemento 1), pp. 7-18.
- MINAYO, M. C. S., SOUZA, E. R., MALAQUIAS, J. V., *et al.*, 2003, “Análise da Morbidade Hospitalar por Lesões e Envenenamentos no Brasil, 2000”. In: Minayo, M.C.S., Souza, E.R. (Org.), *Violência sob o Olhar da Saúde: A Infrapolítica da Contemporaneidade Brasileira*. 1 ed., capítulo 4, Rio de Janeiro, Brasil, Editora FIOCRUZ.
- NETER, J., KUTNER, M. H., NACHTSHEIM, C. J., *et al.*, 1996, *Applied Linear Statistical Models*. 4 ed. New York, McGraw-Hill.

- NUNES, A., SANTOS, J. R. S., BARATA, R. B. *et al.*, 2002, *Medindo as Desigualdades em Saúde no Brasil: Uma Proposta de Monitoramento*. 1 ed. Brasília, Organização Pan-Americana de Saúde, Instituto de Pesquisa Econômica Aplicada.
- OLMSTEAD, T., 2001, “Freeway Management Systems and Motor Vehicle Crashes: A Case Study of Phoenix, Arizona”, *Accident Analysis and Prevention*, v. 33, pp. 433-447.
- ORGANIZAÇÃO MUNDIAL DA SAÚDE, 1996, *Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde. 10ª Revisão*. Tradução: Centro Colaborador da OMS para a Classificação de Doenças em Português. 3 ed. São Paulo, Editora da Universidade de São Paulo.
- PAIM, J. S., COSTA, M. C. N., MASCARENHAS, J. C. S., *et al.*, 1999, “Distribuição Espacial da Violência: Mortalidade por Causas Externas em Salvador (Bahia), Brasil”, *Revista Panamericana de Salud Publica / American Journal of Public Health*, v. 6, n. 5, pp. 321-332.
- PETCH, R. O., HENSON, R. R., 2000, “Child Road Safety in the Urban Environment”, *Journal of Transport Geography*, v. 8, pp. 197-211.
- PINHEIRO, P., 1994, “A Violência do Rio às Portas da Emergência”, *Cadernos de Saúde Pública*, v. 10 (suplemento 1), pp. 223-225.
- PORTUGAL, L. S., SANTOS, M. P. S., 1991, “Trânsito Urbano: a Violência e o seu Contexto Político”, *Revista de Administração Pública*, v. 25, n. 3, pp. 185-197.

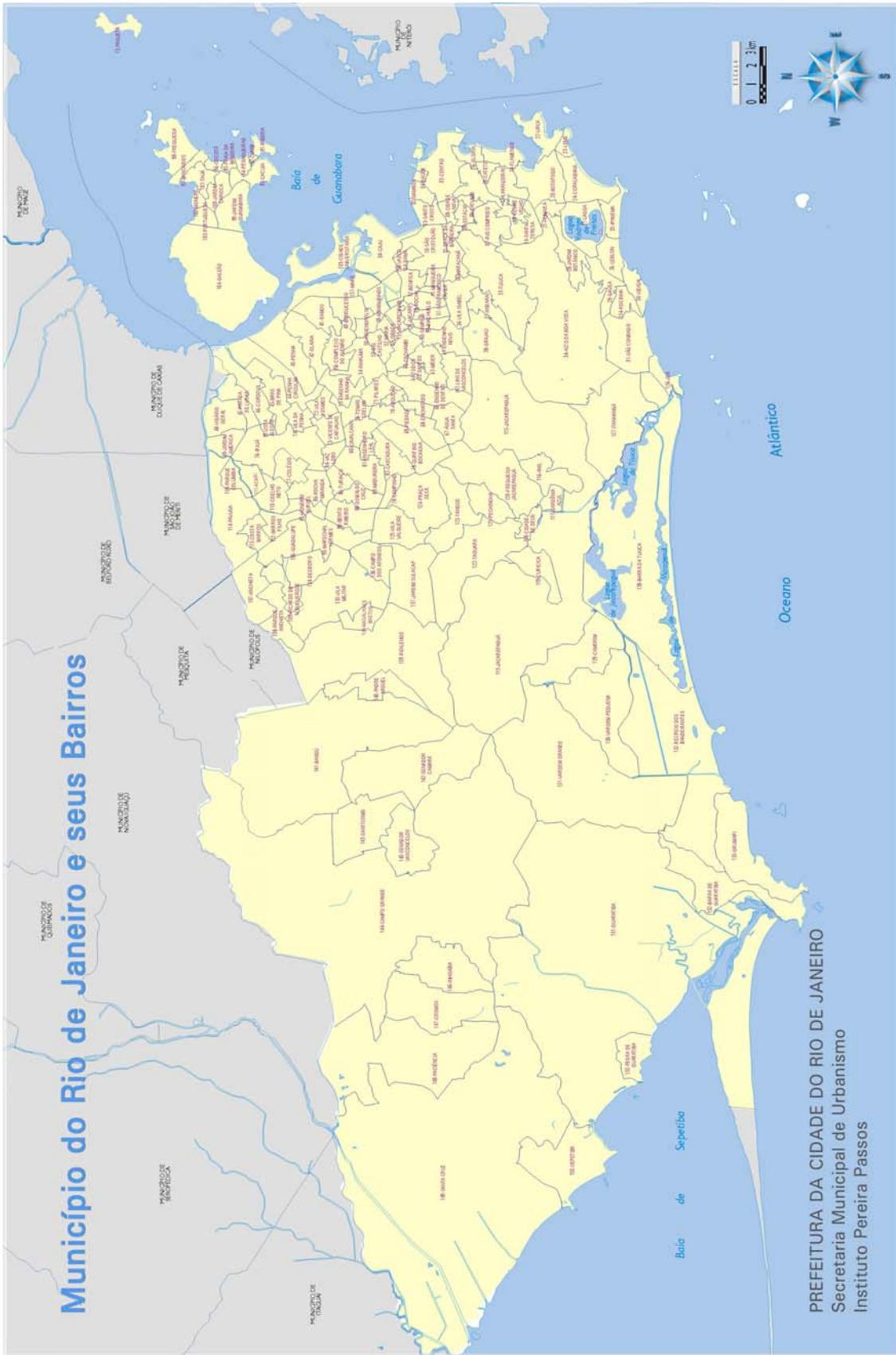
- REICHENHEIM, M., WERNECK, G., 1994, “Anos Potenciais de Vida Perdidos no Rio de Janeiro, 1990. As Mortes Violentas em Questão”, *Cadernos de Saúde Pública*, v. 10 (suplemento 1), pp.188-198.
- RIPLEY, B. D., 1981, *Spatial Statistics*. 1 ed. New York, John Wiley & Sons.
- RODRIGUES, E. M. S., 1995, *Morbidade Hospitalar por Lesões e Envenenamentos, Ribeirão Preto, 1990, em Procedentes do Município*. Tese de D.Sc., Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, Ribeirão Preto, SP, Brasil.
- ROSENBERG, M. L., CARROLL, P. W., POWELL, K. L., 1992, “Let’s Be Clear: Violence is a Public Health Problem”, *JAMA*, v. 267, n. 22, pp. 3071-3072.
- SCALASSARA, M. B., SOUZA, R. K. T., SOARES, D. F. P. P., 1998, “Características da Mortalidade por Acidentes de Trânsito em Localidade da Região Sul do Brasil”, *Revista de Saúde Pública / Journal of Public Health*, v. 32, n. 2, pp. 472-478.
- SCHRAMM, J. M. A., OLIVEIRA, A. F., LEITE, I. C., VALENTE, J. G., GADELHA, A. M. J., PORTELA, M. C., CAMPOS, M. R., 2004, “Transição Epidemiológica e o Estudo da Carga de Doença no Brasil”, *Ciência & Saúde Coletiva*, v. 9, n. 4, pp.897-908.
- Secretaria Municipal de Saúde do Rio de Janeiro, 1996, *Saúde em Foco*, v. 13, pp. 7-9.
- SHANKAR, V., MANNERING, F., BARFIELD, W., 1995, “Effect of Roadway Geometrics and Environmental Factors on Rural Freeway Accident Frequencies”, *Accident Analysis and Prevention*, v. 27, n. 3, pp. 371-389.

- SILVA, C. M. F. P., 2000, *Estudo da Mortalidade por Violência em Idosos, no Brasil e nas Nove Capitais de Regiões Metropolitanas, no Período de 1980 a 1996*. Centro Latino-Americano de Estudos de Violência e Saúde Jorge Careli (CLAVES), Escola Nacional de Saúde Pública (ENSP), Fundação Oswaldo Cruz (FIOCRUZ).
- SIMÕES, C. C. S, 2002, *Perfis de Saúde e de Mortalidade no Brasil: Uma Análise de seus Condicionantes em Grupos Populacionais Específicos*. 2 ed. Brasília, Organização Pan-Americana de Saúde.
- SOUZA, E. R., 1992, *O Reflexo da Violência na Mortalidade da Baixada Fluminense - Brasil*. ENSP/FIOCRUZ.
- SOUZA, E. R., 1993, “Violência Velada e Revelada: Estudo Epidemiológico da Mortalidade por Causas Externas em Duque de Caxias, Rio de Janeiro”, *Cadernos de Saúde Pública*, v. 9, n. 1, pp. 48-64.
- SOUZA, E. R., ASSIS, S. G., SILVA, C. M. F. P., 1997, “Violência no Município do Rio de Janeiro: Áreas de Risco e Tendências da Mortalidade entre Adolescentes de 10 a 19 Anos”, *Revista Panamericana de Salud Pública. Pan American Journal of Public Health*, v. 1, n. 5, pp. 389-398.
- SOUZA, E. R., MINAYO, M. C. S., 1994, “O Impacto da Violência Social na Saúde Pública do Brasil: Década de 80”. In: Minayo, M.C.S. (Org.), *Os Muitos Brasis: Saúde e População na Década de 80*. 1 ed., capítulo 4, São Paulo, Brasil, Hucitec.

SOUZA, E. R., MINAYO, M. C. S., SILVA, C. M. F. P., *et al.*, 2003, “Análise Temporal da Mortalidade por Causas Externas no Brasil: décadas de 80 e 90”. In: Minayo, M.C.S., Souza, E.R. (Org.), *Violência sob o Olhar da Saúde: A Infrapolítica da Contemporaneidade Brasileira*. 1 ed., capítulo 3, Rio de Janeiro, Brasil, Editora FIOCRUZ.

WOOD, G. R., 2005, “Confidence and Prediction Intervals for Generalised Linear Accident Models”, *Accident Analysis and Prevention*, v. 37, pp. 371-389.

ANEXO 1



ANEXO 2

Distribuição do total de acidentes de trânsito com vítimas do Município do Rio de Janeiro, no período de 1998 a 2000, segundo bairro e ano de ocorrência.

Bairro	1998	1999	2000	Total
Abolição	22	27	34	83
Acari	16	14	14	44
Água Santa	22	18	11	51
Alto da Boa Vista	18	21	27	66
Anchieta	69	80	60	209
Andaraí	19	17	10	46
Anil	21	26	35	82
Bancários	18	10	10	38
Bangu	522	510	501	1533
Barra da Tijuca	597	523	650	1770
Barra de Guaratiba	4	5	6	15
Barros Filho	29	16	12	57
Benfica	30	66	58	154
Bento Ribeiro	44	50	54	148
Bonsucesso	386	307	299	992
Botafogo	192	196	156	544
Brás de Pina	60	49	97	206
Cachambi	43	53	44	140
Cacuia	35	33	25	93
Caju	10	33	47	90
Camorim	4	7	6	17
Campinho	62	66	83	211
Campo dos Afonsos	0	2	0	2
Campo Grande	780	846	920	2546
Cascadura	83	71	76	230
Catete	62	43	75	180

Bairro	1998	1999	2000	Total
Catumbi	24	29	33	86
Cavalacanti	7	8	15	30
Centro	848	784	664	2296
Cidade de Deus	59	63	56	178
Cidade Nova	100	109	74	283
Cidade Universitária	17	23	13	53
Cocotá	28	15	13	56
Coelho Neto	57	57	41	155
Colégio	32	38	32	102
Complexo do Alemão	1	0	1	2
Copacabana	285	261	248	794
Cordovil	42	52	34	128
Cosme Velho	7	11	6	24
Cosmos	33	14	27	74
Costa Barros	13	14	16	43
Curicica	53	54	55	162
Del Castilho	56	67	59	182
Deodoro	47	29	42	118
Encantado	24	26	36	86
Engenheiro Leal	0	3	3	6
Engenho da Rainha	15	16	30	61
Engenho de Dentro	67	75	72	214
Engenho Novo	74	102	127	303
Estácio	26	37	38	101
Flamengo	98	98	97	293
Freguesia - ILHA	21	12	8	41
Freguesia - JPA	93	116	89	298
Galeão	94	98	94	286
Gâmboa	0	10	11	21
Gardênia Azul	23	22	13	58

Bairro	1998	1999	2000	Total
Gávea	110	132	168	410
Glória	16	20	11	47
Grajaú	44	59	72	175
Grumari	1	2	4	7
Guadalupe	91	76	72	239
Guaratiba	19	23	13	55
Higienópolis	20	25	13	58
Honório Gurgel	49	59	71	179
Humaitá	24	10	20	54
Inhaúma	68	66	67	201
Inhoaíba	17	30	30	77
Ipanema	89	65	57	211
Irajá	169	163	125	457
Itanhangá	2	4	7	13
Jacaré	46	35	47	128
Jacarepaguá	260	187	411	858
Jacarezinho	11	16	13	40
Jardim América	54	47	41	142
Jardim Botânico	58	47	46	151
Jardim Carioca	20	21	19	60
Jardim Guanabara	44	38	29	111
Jardim Sulacap	54	41	51	146
Joá	3	4	2	9
Lagoa	66	52	80	198
Laranjeiras	37	46	44	127
Leblon	149	158	137	444
Leme	8	27	27	62
Lins de Vasconcelos	38	38	35	111
Madureira	235	211	200	646
Magalhães Bastos	36	26	25	87

Bairro	1998	1999	2000	Total
Mangueira	8	17	15	40
Manguinhos	57	63	38	158
Maracanã	68	78	43	189
Maré	3	4	4	11
Marechal Hermes	126	123	124	373
Maria da Graça	22	19	14	55
Méier	176	198	182	556
Moneró	13	10	12	35
Olaria	88	90	57	235
Oswaldo Cruz	25	37	22	84
Paciência	50	62	52	164
Padre Miguel	79	95	61	235
Paquetá	1	1	1	3
Parada de Lucas	57	72	52	181
Parque Anchieta	9	1	14	24
Pavuna	141	140	131	412
Pechincha	33	38	32	103
Pedra de Guaratiba	9	15	25	49
Penha	201	267	259	727
Penha Circular	6	22	10	38
Piedade	119	99	91	309
Pilares	60	85	61	206
Pitangueiras	4	5	4	13
Portuguesa	35	31	36	102
Praça da Bandeira	100	76	127	303
Praça Seca	74	82	75	231
Praia da Bandeira	0	1	0	1
Quintino Bocaiúva	34	39	41	114
Ramos	90	109	79	278
Realengo	366	371	375	1112

Bairro	1998	1999	2000	Total
Recreio dos Bandeirantes	97	112	124	333
Riachuelo	17	18	21	56
Ribeira	2	4	4	10
Ricardo de Albuquerque	55	62	86	203
Rio Comprido	89	76	62	227
Rocha	13	14	13	40
Rocha Miranda	91	68	53	212
Rocinha	22	11	10	43
Sampaio	12	12	17	41
Santa Cruz	371	337	380	1088
Santa Teresa	15	10	26	51
Santíssimo	49	54	65	168
Santo Cristo	15	27	34	76
São Conrado	59	48	38	145
São Cristóvão	451	392	332	1175
São Francisco Xavier	24	19	15	58
Saúde	1	5	5	11
Senador Camará	47	71	69	187
Senador Vasconcelos	10	5	8	23
Sepetiba	33	41	30	104
Tanque	33	34	31	98
Taquara	178	219	156	553
Tauá	30	26	22	78
Tijuca	261	275	261	797
Todos os Santos	20	28	67	115
Tomas Coelho	18	9	21	48
Turiaçu	7	15	6	28
Urca	6	10	11	27
Vargem Grande	9	19	28	56
Vargem Pequena	10	20	25	55

Bairro	1998	1999	2000	Total
Vaz Lobo	12	25	22	59
Vicente de Carvalho	59	81	99	239
Vidigal	4	4	6	14
Vigário Geral	36	18	19	73
Vila da Penha	72	64	48	184
Vila Isabel	122	135	128	385
Vila Kosmos	8	8	11	27
Vila Militar	17	12	5	34
Vila Valqueire	61	43	41	145
Vista Alegre	22	9	20	51
Zumbi	11	18	7	36

ANEXO 3

Distribuição do total de acidentes de trânsito com vítimas do Município do Rio de Janeiro, no período de 1998 a 2000, segundo bairro de ocorrência e tipo de evento.

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Abolição	83	40	43
Acari	44	15	29
Água Santa	51	35	16
Alto da Boa Vista	66	48	18
Anchieta	209	90	120
Andaraí	46	17	29
Anil	82	36	46
Bancários	38	17	21
Bangu	1533	595	944
Barra da Tijuca	1770	1149	626
Barra de Guaratiba	15	6	9
Barros Filho	57	24	33
Benfica	154	65	89
Bento Ribeiro	148	67	82
Bonsucesso	992	466	531
Botafogo	544	232	314
Brás de Pina	206	101	106
Cachambi	140	64	76
Cacua	93	36	57
Caju	90	42	49
Camorim	17	4	13
Campinho	211	90	122
Campo dos Afonsos	2	2	0
Campo Grande	2546	1121	1437
Cascadura	230	75	156

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Catete	180	53	128
Catumbi	86	30	56
Cavalacanti	30	15	15
Centro	2296	635	1665
Cidade de Deus	178	66	113
Cidade Nova	283	127	156
Cidade Universitária	53	31	22
Cocotá	56	26	30
Coelho Neto	155	69	86
Colégio	102	52	50
Complexo do Alemão	2	0	2
Copacabana	794	286	513
Cordovil	128	60	69
Cosme Velho	24	16	8
Cosmos	74	27	48
Costa Barros	43	12	31
Curicica	162	59	103
Del Castilho	182	61	121
Deodoro	118	71	47
Encantado	86	50	36
Engenheiro Leal	6	1	5
Engenho da Rainha	61	20	41
Engenho de Dentro	214	128	86
Engenho Novo	303	140	163
Estácio	101	32	69
Flamengo	293	99	195
Freguesia - ILHA	41	18	23
Freguesia - JPA	298	153	145
Galeão	286	191	95

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Gâmbua	21	6	15
Gardênia Azul	58	20	38
Gávea	410	219	193
Glória	47	14	33
Grajaú	175	78	97
Grumari	7	6	1
Guadalupe	239	110	129
Guaratiba	55	18	37
Higienópolis	58	29	29
Honório Gurgel	179	88	92
Humaitá	54	25	29
Inhaúma	201	93	109
Inhoaíba	77	19	58
Ipanema	211	112	99
Irajá	457	226	233
Itanhangá	13	7	6
Jacaré	128	43	85
Jacarepaguá	858	419	442
Jacarezinho	40	5	35
Jardim América	142	80	62
Jardim Botânico	151	82	70
Jardim Carioca	60	37	23
Jardim Guanabara	111	65	46
Jardim Sulacap	146	92	54
Joá	9	7	2
Lagoa	198	130	69
Laranjeiras	127	51	77
Leblon	444	256	191
Leme	62	15	47

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Lins de Vasconcelos	111	36	75
Madureira	646	232	416
Magalhães Bastos	87	33	54
Mangureira	40	8	32
Manguinhos	158	60	99
Maracanã	189	74	115
Maré	11	0	11
Marechal Hermes	373	140	235
Maria da Graça	55	28	27
Méier	556	191	366
Moneró	35	16	19
Olaria	235	119	118
Oswaldo Cruz	84	42	42
Paciência	164	45	120
Padre Miguel	235	77	159
Paquetá	3	1	2
Parada de Lucas	181	125	56
Parque Anchieta	24	5	19
Pavuna	412	163	249
Pechincha	103	53	50
Pedra de Guaratiba	49	20	29
Penha	727	419	310
Penha Circular	38	18	20
Piedade	309	131	178
Pilares	206	79	127
Pitangueiras	13	6	7
Portuguesa	102	38	64
Praça da Bandeira	303	102	201
Praça Seca	231	107	124

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Praia da Bandeira	1	1	0
Quintino Bocaiúva	114	47	67
Ramos	278	139	142
Realengo	1112	492	625
Recreio dos Bandeirantes	333	181	153
Riachuelo	56	24	33
Ribeira	10	6	4
Ricardo de Albuquerque	203	93	111
Rio Comprido	227	89	139
Rocha	40	21	21
Rocha Miranda	212	94	118
Rocinha	43	9	34
Sampaio	41	17	24
Santa Cruz	1088	454	638
Santa Teresa	51	17	34
Santíssimo	168	71	97
Santo Cristo	76	22	54
São Conrado	145	84	62
São Cristóvão	1175	525	651
São Francisco Xavier	58	25	33
Saúde	11	3	8
Senador Camará	187	57	130
Senador Vasconcelos	23	9	14
Sepetiba	104	48	56
Tanque	98	45	55
Taquara	553	295	259
Tauá	78	38	41
Tijuca	797	323	475
Todos os Santos	115	50	65

Bairro	Acidentes de Trânsito	Colisões	Atropelamentos
Tomas Coelho	48	25	23
Turiaçu	28	11	17
Urca	27	8	19
Vargem Grande	56	21	35
Vargem Pequena	55	26	29
Vaz Lobo	59	25	34
Vicente de Carvalho	239	118	122
Vidigal	14	4	10
Vigário Geral	73	28	45
Vila da Penha	184	108	79
Vila Isabel	385	165	223
Vila Kosmos	27	16	11
Vila Militar	34	21	13
Vila Valqueire	145	62	83
Vista Alegre	51	26	25
Zumbi	36	20	16