

# **“APLICAÇÃO DA TÉCNICA DE ANÁLISE DE DADOS ESPACIAIS DO TIPO ORIGEM-DESTINO APLICADOS A DADOS GEOREFERENCIADOS DE ACIDENTES DE TRÂNSITO OCORRIDOS EM NATAL/RN”**

Eduardo Henrique Silveira de Araújo, Departamento de Estatística/UFRN [ehsa@ccet.ufrn.br](mailto:ehsa@ccet.ufrn.br); Flávia Ferreira Batista, curso de Estatística/UFRN; Marcilio R. de Oliveira, curso de Estatística/UFRN – [marcilioju@gmail.com](mailto:marcilioju@gmail.com) ; Franciné dos Santos Pessoa, Departamento de Estatística/UFRN – [fsp@ccet.ufrn.br](mailto:fsp@ccet.ufrn.br).

## **RESUMO**

O presente estudo tem como principal foco a análise espacial pontual do tipo Origem-Destino aplicada a dados de acidentes de trânsito georeferenciados ocorridos na malha viária de Natal-RN no 2º trimestre do ano de 2008. Nesse estudo o “Destino” se refere à localização georeferenciada dos acidentes de trânsito e a “Origem” se refere a localização georeferenciada dos endereços dos veículos envolvidos nos acidentes da amostra analisada. O estudo teve o objetivo de Identificar na distribuição espacial dos acidentes ocorridos malha viária de Natal e do local dos dois endereços dos veículos envolvidos nos acidentes, regiões que poderiam apresentar formação de aglomerados de acidentes. Através da função M foi possível testar a hipótese de associação da distribuição dos processos pontuais de Origem e Destino. Foram utilizados os softwares ArcView e TerraView para a construção dos mapas temáticos da distribuição espacial georeferenciada dos acidentes e dos endereços dos veículos. Foi elaborado o algoritmo para estimar a função M com simulações dos valores de K1 e K2 referente aos acidentes e aos endereços. O resultado da aplicação desta função indicou que existe associação entre os padrões pontuais.

## **1) OBJETIVOS**

Identificar se existe correlação entre a distribuição espacial dos acidentes de trânsito georeferenciados ocorridos na malha viária do município de Natal nos meses do primeiro semestre de 2008 e a distribuição espacial dos endereços georeferenciados dos veículos envolvidos através da técnica de análise estatística espacial do tipo Origem-Destino.

## **2) LOCALIZAÇÃO DA ÁREA ESTUDADA**

Conhecida como a "Cidade do Sol", cidade localizada numa região chamada "esquina do continente", capital do estado brasileiro do Rio Grande do Norte, pertencente à Região Metropolitana de Natal. Situa-se numa espécie de triângulo natural com um vértice para o norte, que é banhado de um lado pelo Rio Potengi e de outro pelo Oceano Atlântico. O número de acidentes de trânsito mensal ocorridos em Natal é um fator preocupante para as autoridades gerenciadoras do trânsito do município. A média mensal de acidentes de trânsito ocorridos é 720. Embora seja uma das menores entre as capitais do Brasil, esta média é preocupante para as autoridades dos órgãos gerenciadores. Alguns estudos científicos sobre segurança no trânsito urbana foram desenvolvidos por profissionais e consultores dos órgãos gerenciadores do trânsito em Natal, principalmente para subsidiar políticas de prevenção de acidentes em cruzamentos de ruas e avenidas da malha viária de Natal. Esses estudos científicos foram realizados através da modelagem estatística temporal e espacial. Os estudos sobre a distribuição espacial georeferenciadas dos acidentes, tem produzidos resultados excelentes e foram utilizados pelo Setor de Estatística do DETRAN/RN para subsidiar políticas de prevenção de acidentes na malha viária de Natal.

### 3) METODOLOGIA DA PESQUISA

A metodologia aplicada para o desenvolvimento da pesquisa visando atingir os objetivos apresentados no segundo capítulo desta monografia é composta por várias etapas. Na primeira etapa estruturou-se o banco de dados dos acidentes de trânsito de veículos obtidos no setor de Estatística do DETRAN-RN. Na segunda etapa com base no banco de dados de acidentes de trânsito foram selecionadas algumas variáveis de interesse para a análise estatística espacial: tipo do acidente, nº de veículos envolvidos no acidente, tipo do veículo, dados do proprietário/condutor do veículo. Na terceira etapa estruturado o banco de dados com o georeferenciamento de cada um dos acidentes no ambiente do software **ArcView 3.2**. Na quarta etapa selecionou-se no banco de dados os acidentes ocorridos no período dos primeiros seis meses do ano de 1998 que envolveram dois veículos do tipo automóvel e que eram de origem no município de Natal. Na quinta etapa foram identificados os endereços de origem dos dois veículos envolvidos nos acidentes. Foi criado um novo banco de dados com os endereços georeferenciados dos dois veículos. Na sexta etapa foi selecionado no banco de dados dos acidentes georeferenciados aqueles que envolviam os dois veículos do tipo automóvel e com endereço em Natal. Na oitava etapa foi feita no ambiente do **TerraView 3.30** a construção de mapas temáticos com a distribuição espacial pontual georeferenciada dos acidentes e dos endereços dos veículos envolvidos. Na nona etapa foi realizada todas as análises estatísticas espaciais dos acidentes e dos endereços dos dois veículos e construídos os mapas do alisamento de Kernel. Também nesta etapa foram estimados os valores das dependências espaciais e a estimativa da função M para a ligação dos processos de distribuição espacial dos eventos e dos endereços. Na décima etapa foi feita a conclusão do estudo.

### 5) HIPÓTESES TESTADAS NA ANÁLISE DE ASSOCIAÇÃO ESPACIAL DE ORIGEM-DESTINO.

Para atender aos objetivos do trabalho, foram elaboradas as hipóteses descritas abaixo que retratam a associação espacial entre os eventos que representam a origem (endereços dos veículos) e o destino (local do acidente).

#### Hipóteses Científicas:

**H<sub>0</sub>**: Não Existe associação entre a distribuição espacial dos acidentes na malha viária de Natal (destino) e a distribuição espacial da localização (origem) dos veículos envolvidos e **H<sub>A</sub>**: Caso Contrário.

#### Hipóteses Estatísticas:

**H<sub>0</sub>**:  $M_{obs} \neq M_{est}$  e **H<sub>A</sub>**:  $M_{obs} = M_{est}$ , com  $M_{obs}$  matriz de valores de M observados e  $M_{est}$  matriz de valores de M estimado.

### 6) DEPENDÊNCIA ESPACIAL E ALISAMENTO DE KERNEL

Para testar as hipóteses de associação espacial entre eventos, inicialmente foram realizadas análises da dependência espacial entre os eventos que representam o padrão pontual do conjunto de dados das localizações  $A = (\rho_1, \rho_2, \dots, \rho_n)$  dos acidentes de trânsito e também entre os eventos que representam o padrão pontual do conjunto de dados das localizações  $E_1 = (\rho_1, \rho_2, \dots, \rho_n)$  dos endereços do primeiro veículo e  $E_2 = (\rho_1, \rho_2, \dots, \rho_n)$  dos endereços do segundo veículo envolvido no acidente. Todos os pontos foram representados geometricamente por um par ordenado  $(x, y)$  de

coordenadas espaciais. A Dependência Espacial está associada a duas hipóteses: hipótese nula: a distribuição espacial dos dados obtidos para a variável em análise é aleatória, ou seja, existe independência espacial e a hipótese alternativa: existência de dependência espacial. Para testar as hipóteses, foi realizada para todos os três padrões pontuais análise da autocorrelação espacial, que possibilita identificar qual tipo de relação existe entre os eventos pontuais ou a intensidade entre eventos. Para identificar se existe a autocorrelação foram estimados os valores da função K de Kernel, estimador de intensidade dos eventos (acidentes - destino ou endereços - origem). A função K apresenta grande poder para detectar em diferentes escalas, se os eventos estão distribuídos ao acaso, se eles se encontram em padrão de regularidade, ou se existem conglomerados específicos, proporciona uma descrição efetiva da dependência espacial dentro de um padrão pontual em determinada área.

$$\hat{K}(h) = \frac{A}{n^2} \sum_i^n \sum_{j, j \neq i}^n \frac{I_h(d_{ij})}{w_{ij}}$$

Onde: **A** é a área da região, **n** é o número de eventos observados, **I<sub>h</sub>(d<sub>ij</sub>)** é uma função indicatriz cujo valor é 1 se (d<sub>ij</sub>) ≤ h e 0 em caso contrário e **w<sub>ij</sub>** é a proporção da circunferência do círculo centrado no evento i que está dentro da região (correção devido ao efeito de borda).

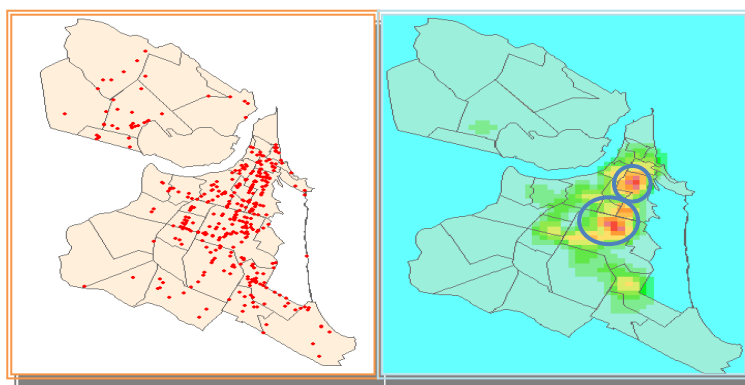


Figura 01 – Distribuição espacial dos acidentes de trânsito com dois veículos na cidade de Natal/RN, 2008 e Mapa de Kernel.

## 7) APLICAÇÃO DA TÉCNICA ORIGEM / DESTINO E ESTIMAÇÃO DA FUNÇÃO M

A técnica de análise espacial do tipo Origem-Destino tem o objetivo de verificar a existência de associação (ligação) entre dois processos de padrões pontuais de eventos ocorridos em uma área de superfície terrestre  $A \subset R_2$  georeferenciada. Como já citado anteriormente, neste trabalho a origem foram identificadas como sendo os endereços georeferenciados dos dois veículos envolvidos no acidente e o destino como sendo a localização do acidente georeferenciado na malha viária de Natal. Neste trabalho o objetivo é verificar se existe associação entre a distribuição espacial dos acidentes ocorridas na malha viária de Natal e os endereços dos veículos envolvidos. Assim por exemplo, pela técnica é possível se verificar se um ponto (evento) georeferenciado de origem (endereço) tem ligação direta com determinado ponto de destino (local do acidente). Para inferir a força de ligação entre dois processos utilizamos a função M criada pelo Prof. Renato Assunção onde a partir dela é possível verificar se, em geral, pontos de origem que ocorrem próximos entre si geram pontos de destino que também são próximos entre si. Segundo ASSUNÇÃO (2006) a função M pode ser interpretada como, um processo pontual bivariado ( $N_1, N_2$ ) em um polígono

formado pelos eventos de  $x_1, \dots, x_n$  de  $N_1$  e  $y_1, \dots, y_n$  de  $N_2$  onde  $x_1 = (x_{1i}, x_{2i})$  e  $y_1 = (y_{1i}, y_{2i})$ . Os processos  $N_1$  e  $N_2$  são ligados de forma que todo evento em  $N_1$  causa um evento em  $N_2$  a ser observado dentro de  $A$ . Denomina-se  $N_1$  e  $N_2$  respectivamente por processo de origem e de destino, como podemos observar na figura 10. Nesse estudo, por exemplo, temos  $n$  pares de locais endereços dos veículos em  $x_i$  e locais de acidentes em  $y_i$ ,  $i = 1, \dots, n$ .

A função  $M(k_1, k_2)$ . pode ser estimada como o número médio observado dos eventos na realização que pertençam à vizinhança (cujo tamanho é definido por  $k_1$  e  $k_2$ ) do evento arbitrário tanto na origem quanto no destino. Fixando-se os valores de  $k_1$  e  $k_2$ , para cada evento de origem  $x_i$ .

A estimativa da função é dada por:

$$\tilde{M}(k_1; k_2) = \frac{\sum_{i=1}^n m_i}{n}$$

Sendo  $m_i$ , o número de eventos que estão entre os  $k_1$  vizinhos mais próximos de  $x_i$  cujos destinos estão entre os  $k_2$  eventos mais próximos de  $y_i$ , destino de  $x_i$ .

Tanto os  $k_1$  pontos de origem quanto os demais possuem igual probabilidade de estarem ligados a algum daqueles  $k_2$  pontos de destino vizinhos mais próximos do ponto de destino em questão. Assim a distribuição de probabilidade dos pontos  $K_1$  e  $K_2$  é hipergeométrica  $(n-1, k_1, k_2)$ . Sob a hipótese de aleatoriedade da ligação entre os processos o valor esperado do estimador é:

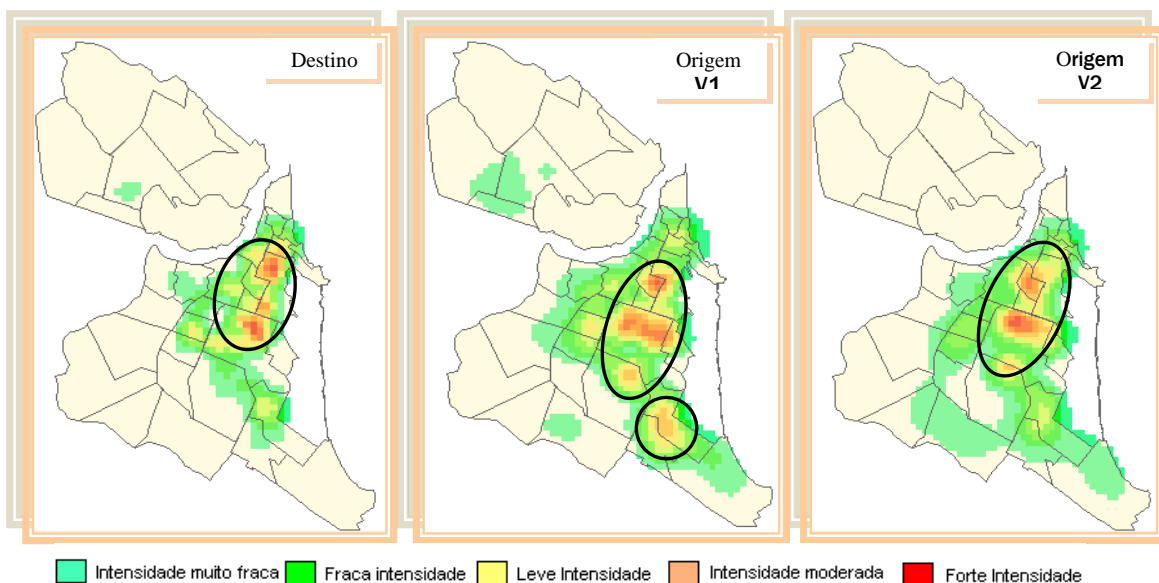
$$E[\tilde{M}(k_1; k_2)] = \frac{k_1 * k_2}{n - 1} = \bar{m}$$

Para a realização do teste, mantêm-se as posições de ambas as realizações, mas permutam-se as ligações entre os processos e para cada uma dessas permutações computa-se o valor de  $M(k_1, k_2)$ .

## RESULTADOS DA APLICAÇÃO DO MÉTODO

### 1) CONSTRUÇÃO DOS MAPAS DE KERNEL

Na figura abaixo pode ser visto os mapas de Kernel do destino e das origens. Verifica-se no mapa de destino (local do acidente) a formação de aglomerados de maior intensidade aproximadamente semelhante ao mapa da origem (endereço) dos veículos V2. Podemos inferir que a maior parte dos veículos envolvidos nestes locais de acidentes origina-se dessas áreas e de áreas próximas. Para os veículos V1 verificam-se outros aglomerados.



**Figura 02** - Mapas temáticos da distribuição espacial dos Acidentes de trânsito e das origens dos veículos envolvidos no 1º semestre de 2008.

## 2) ESTIMATIVA DA FUNÇÃO M E TESTE DE ALEATORIZAÇÃO

Para o cálculo da função M foram fixado do total dos acidentes valores de  $k_1$  e  $k_2$ , iguais a 10, 20, 30, 40, 50, 60. Os valores observados para a origem do veículo 1 encontram-se dispostos na Tabela 01, enquanto na Tabela 02 são apresentados os valores esperados sobre hipótese de aleatoriedade na ligação entre os processos de origem-destino. Como o objetivo é verificar se a distribuição pontual do processo de origem tem associação (ligação) com a distribuição pontual do processo de destino e vice-versa, a hipótese a ser testada é:

$H_0$ : A ligação entre os processos de Origem-Destino é aleatória, contra

$H_A$ : Existe um padrão de ligação entre os processos (aglomerado, regular).

Tabela 01 - Mobs

K1 \ K2	10	20	30	40	50	60
10	0.25	0.58	0.83	1.13	1.39	1.68
20	0.48	1.08	1.55	2.14	2.68	3.26
30	0.81	1.66	2.41	3.33	4.11	4.95
40	1.09	2.16	3.19	4.39	5.50	6.65
50	1.38	2.74	4.00	5.47	6.84	8.27
60	1.63	3.23	4.73	6.49	8.18	9.84

Tabela 02 - Mest

K1 \ K2	10	20	30	40	50	60
10	0.24	0.48	0.72	0.96	1.20	1.44
20	0.48	0.96	1.44	1.92	2.40	2.88
30	0.72	1.44	2.16	2.88	3.60	4.32
40	0.96	1.92	2.88	3.84	4.80	5.76
50	1.20	2.40	3.60	4.80	6.00	7.19
60	1.44	2.88	4.32	5.76	7.19	8.63

Tabela 03-Mobs

K1 \ K2	10	20	30	40	50	60
10	0.37	0.68	0.95	1.24	1.48	1.71
20	0.72	1.35	1.88	2.42	2.90	3.35
30	1.00	1.90	2.70	3.45	4.16	4.91
40	1.26	2.47	3.47	4.49	5.46	6.50
50	1.55	2.99	4.27	5.49	6.75	8.06
60	1.78	3.49	5.03	6.52	8.06	9.67

Tabela 04- Mest

K1 \ K2	10	20	30	40	50	60
10	0.24	0.48	0.72	0.96	1.20	1.44
20	0.48	0.96	1.44	1.92	2.40	2.88
30	0.72	1.44	2.16	2.88	3.60	4.32
40	0.96	1.92	2.88	3.84	4.80	5.76
50	1.20	2.40	3.60	4.80	6.00	7.19
60	1.44	2.88	4.32	5.76	7.19	8.63

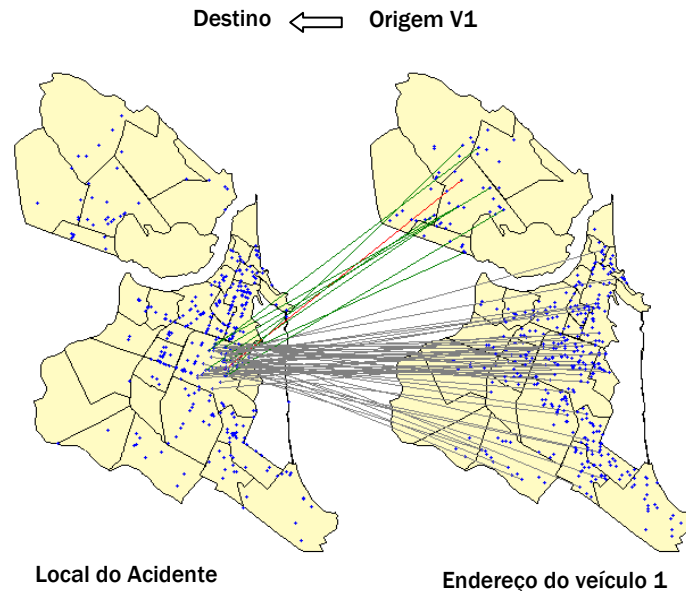
**Figura 04** - Valores observados (tabela 01) e estimados (tabela 02) para a estimativa da função M para dados de origem veículo 1 e local de acidente em Natal sob a suposição de ligação aleatória entre os padrões pontuais. Valores observados (tabela 03) e estimados (tabela 04) para a estimativa da função M para dados de origem veículo 2 e local de acidente em Natal sob a suposição de ligação aleatória entre os padrões pontuais.

## 3) CONCLUSÃO

1) A partir do teste de comparação entre as tabelas percebe-se que os valores observados para a estatística M encontram-se aproximados de seus valores esperados sob hipótese nula de aleatoriedade na ligação entre os padrões. Pelo teste Qui-Quadrado verifica-se se existe dependência entre a origem do veículo 1 e o destino, a estatística do teste foi  $\chi^2 = 1,99$  com 36 graus de liberdade e p-valor  $\alpha > 0,05$ , portanto não rejeitamos assim a hipótese nula. Assim conclui-se que existe um forte indício de que a ligação entre a localização dos veículos V1 (primeiro veículo envolvido) e o local do acidente seja aleatória.

2) A partir do teste de comparação entre as tabelas relacionadas à origem do veículo 2 e o destino percebe-se que os valores observados para a estatística M

encontram-se consideravelmente aproximados de seus valores esperados sob hipótese nula de aleatoriedade na ligação entre os padrões. Pelo teste Qui-Quadrado verifica-se que existe dependência entre a origem do veículo 2 e o destino. A estatística do teste foi  $\chi^2 = 3,84$  com 36 graus de liberdade e p-valor  $\alpha > 0,05$ , portanto não rejeitamos assim a hipótese nula. Assim conclui-se que existe um forte indício de que a ligação entre a localização dos veículos V2 (segundo veículo envolvido) e o local do acidente seja aleatória. Como exemplo, os mapas da figura 05 abaixo retratam bem a primeira conclusão.



**Figura 05** - Mapas da função M com  $k_1 = 60$  e  $k_2 = 60$  para os dados de local do acidente de trânsito e local de endereço dos veículos referente ao acidente 226.

## BIBLIOGRAFIA

- ASSUNÇÃO, Renato M.; LOPES, Danilo L. Medida de correlação entre Padrões pontuais de origem-destino. VIII Brazilian Symposium on GeoInformatics, Campos do Jordão, Brasil, November 19-22, 2006, INPE, p. 315-320.
- ASSUNÇÃO, Renato M. Estatística espacial com aplicações em epidemiologia. ABE – Associação Brasileira de Estatística, 31ª Reunião Regional: Natal, 1999.
- DRUCK, S.; CARVALHO, M.S.; CÂMARA, G.; MONTEIRO, A.V.M. (eds). Análise Espacial de Dados Geográficos. Brasília, EMBRAPA, 2004. Capítulo 1 - Análise Espacial e Geoprocessamento e Capítulo 2- Análise Espacial de Eventos.
- NOBRE, Flávio Fonseca. Introdução à análise de dados espaciais. Programa de Engenharia Biomédica, COPPE/UFRJ: 2001.
- SOUZA, Cleuneide Rodrigues de; Monografia de Graduação - Análise Estatística Espacial dos Acidentes de Trânsito Georeferenciados no Município de Natal/RN no ano de 2007. Natal, UFRN, 2008.