

Método Scan Espaço-Temporal: uma avaliação dos modelos Poisson e Permutação Espaço-Tempo

Sadraque Enéas de Figueirêdo Lucena, Ronei Marcos de Moraes

Laboratório de Estatística Aplicada ao Processamento de Imagens e Geoprocessamento

Departamento de Estatística – Universidade Federal da Paraíba

INTRODUÇÃO

Estudos de detecção de conglomerados são imprescindíveis para identificação de regiões geográficas cujo risco em relação a um determinado evento, seja uma doença ou uma epidemia, em suas diversas formas é elevado [1]. Na literatura existe uma vasta quantidade de métodos de detecção de conglomerados, de modo que estes são classificados de acordo com as características e as hipóteses feitas sobre os conglomerados [2]. Em ambos os métodos a hipótese de nula considerada supõe que não há conglomerados em toda região, ou seja, o risco de ocorrência do evento é completamente aleatório e proporcional ao tamanho populacional de cada localidade [3].

Dentre os métodos de detecção de conglomerados, o método Scan (ou método Scan) tem sido amplamente usado, pois além de delinear regiões críticas por meio de algoritmos computacionais gráficos, é atribuída significância à estatística associada via simulação de Monte Carlo [4]. Esta metodologia pode ser aplicada à dados espaciais, temporais ou espaço-temporais, para vários modelos de probabilidade [5]. Dentre os modelos de probabilidade ganham destaque o Poisson e o Permutação Espaço-Tempo quando se trata de dados de contagem [6] [7].

Quanto se tem mais de um método passível de aplicação sobre um conjunto de dados, muitas vezes convém avaliar o desempenho dos métodos para identificar qual deles apresenta melhores resultados. Sob esta ótica, o presente trabalho se dispôs a avaliar o desempenho do Método Scan Espaço-Temporal utilizando os modelos Poisson e Permutação Espaço-Tempo. Para tanto foram realizadas simulações utilizando uma região hipotética.

MATERIAIS E MÉTODOS

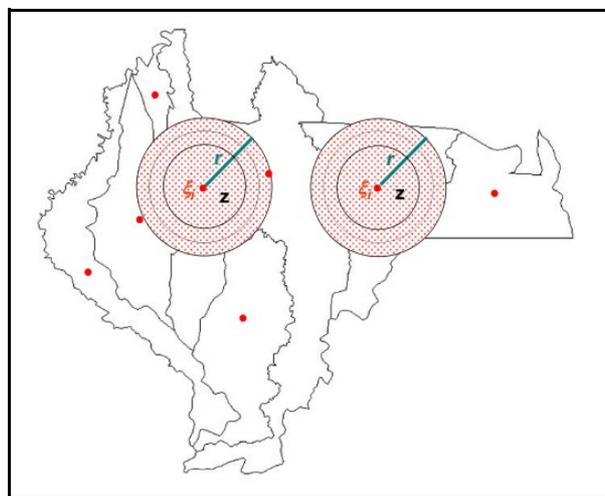
Para realizar as simulações foi elaborada uma região hipotética particionada em 85 sub-áreas. Tendo em vista uma análise do tipo espaço-temporal, foi determinado que cada simulação corresponderia a cinco períodos de tempo distintos e consecutivos. Os dados obtidos das simulações foram associados a cada sub-área da região hipotética, de modo a introduzir o conceito espacial aos dados. Assim, foram realizadas 50 simulações, cada uma com cinco períodos, de modo a permitir conclusões expressivas sobre os resultados. Para gerar os dados foi utilizada a distribuição qui-quadrado com 1 grau de liberdade, em seguida os valores foram truncados, permitindo assim a obtenção de dados do tipo discreto e com assimetria à direita. Essa característica é comumente

encontrada em dados de contagem de estudos epidemiológicos [8]. Para cada sub-área também foi atribuída uma população específica que variou em cada um dos cinco períodos, simulando um crescimento populacional.

Embora tenha sido realizada uma análise espaço-temporal no presente artigo, faz-se necessária a introdução da abordagem puramente espacial. A seguir são descritos os métodos utilizados e apresentada a formulação dos modelos Poisson e Permutação Espaço-Tempo.

Scan Espacial

Seja uma área de estudo particionada em sub-regiões menores. A cada sub-região é associado um centroide ξ_i , que contém os respectivos número de indivíduos com determinada doença e o tamanho da população na sub-região. O método proposto por Kulldorff e Nagarwalla [7] se baseia em círculos posicionados sobre cada centroide ξ_i , cujo tamanho do raio r pode ser aumentado de modo a envolver centroides vizinhos. O círculo é aumentado até conter dentro dele um determinado percentual da população total. À medida que os círculos aumentam é feito um teste de significância para verificar se a região delimitada pela circunferência corresponde à um conglomerado. O teste é baseado no método de máxima verossimilhança, podendo ser assumidas algumas distribuições de probabilidade para a variável de interesse. A avaliação é realizada por simulações de Monte Carlo [8]. É possível também utilizar outras forma geométricas além da circunferência, como uma elipse ou outro formato arbitrário [4]. A Figura 1 apresenta de forma visual um exemplo hipotético da varredura espacial do método Scan circular.



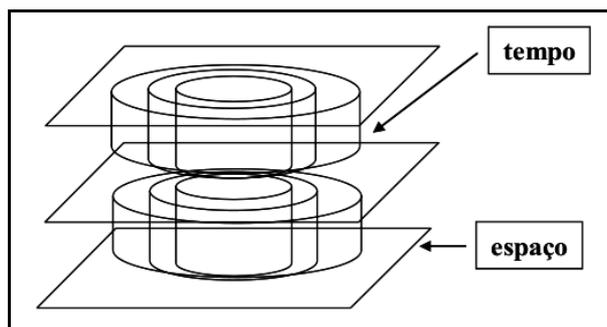
FONTE: Adaptado de BALIEIRO (2008)

Figura 1. Exemplo hipotético da varredura espacial da método Scan

Scan Espaço-Temporal

Além da análise espacial, o método Scan permite incorporar o fator temporal no

estudo. Neste caso é de interesse identificar conglomerados de doenças que venham a ocorrer no espaço e no tempo simultaneamente. O procedimento é análogo ao método espacial, sendo que neste caso, ao invés de círculos, são utilizados cilindros, cuja base representa a o espaço geográfico e a altura representa o tempo. Embora o procedimento permita a base e a altura variarem continuamente de acordo com a varredura do método no espaço e no tempo, não há alterações em relação à razão de verossimilhança e a significância do teste [10]. A Figura 2 apresenta de maneira visual o procedimento da varredura espaço-temporal.



FONTE: Adaptado de BALIEIRO (2008)

Figura 2. Exemplo hipotético da varredura espaço-temporal da método Scan

No que diz respeito ao caráter temporal de um estudo com a método Scan, duas abordagens podem ser utilizadas: a análise retrospectiva e a prospectiva. A análise retrospectiva tem por objetivo detectar tanto conglomerados ainda ativos no fim do período de estudo, quanto conglomerados que deixaram de existir antes do fim do período [5]. A análise prospectiva busca detectar apenas conglomerados que não deixaram de ocorrer até o final do estudo. Nesta segunda abordagem, o processo de detecção é iniciado varrendo-se o primeiro intervalo de tempo (dia, mês, ano, etc.) e, em seguida, os intervalos subsequentes levam em consideração as regiões identificadas anteriormente [11].

Para realizar as simulações dos dados foi utilizado o *software* R [12]. A aplicação do método Scan foi realizada no *software* SaTScan [5]. Tendo em vista que a saída do SaTScan não apresenta os resultados em mapas, os resultados foram importados para o R para elaboração dos mesmos.

RESULTADOS E DISCUSSÃO

A abordagem utilizada para análise dos mapas se baseou no conceito epidemiológico, o qual espera que os conglomerados identificados não ocorram em regiões muito distintas de onde se evidenciaram os casos. Tendo em vista que o método Scan se baseia no conceito de risco para determinar um conglomerado, foram gerados mapas de risco para cada sub-região. Para obter os mapas de risco foi calculada a incidência de cada sub-área (quociente entre o valor observado na sub-área e a

respectiva população) e a incidência correspondente à área total (quociente entre a soma de todos os valores observados e a soma da população). Em seguida o risco foi obtido pela divisão entre a incidência de cada sub-área pela incidência da área total. Assim, pôde-se identificar as sub-áreas com mais risco de ocorrência do evento.

Para aplicar o método Scan espaço-temporal foi determinado que o conglomerado não deveria ser superior a 50% do tamanho populacional e a significância utilizada para testar a hipótese nula de aleatoriedade foi 0,05. O método de análise empregado foi do tipo retrospectivo, permitindo assim identificar conglomerados que persistiram não só até o fim do período, como também aqueles que deixaram de existir antes do término do período.

Ao comparar os mapas de risco com os mapas de ambos modelos, verificou-se em casos de conglomerados que ocorreram unicamente em período de tempo, o modelo Permutação Espaço-tempo identificou conglomerados maiores que o modelo Poisson. Foi constatado também que na ocorrência de sub-áreas próximas, mas que não eram vizinhas, o modelo Poisson indicou conglomerados apenas na sub-área com maior risco. Por outro lado, o modelo Permutação Espaço-Tempo identificou ambas sub-áreas, além de sub-áreas adjacentes.

Os conglomerados espaço-temporais identificados por ambos os métodos ocorreram em no máximo dois períodos consecutivos. O modelo Poisson tendeu a identificar conglomerados espaço-temporais em mais períodos de tempo consecutivos que o modelo Permutação Espaço-Tempo. Por exemplo, na Figura 3 o modelo Poisson (à esquerda) identificou quatro períodos consecutivos distintos (P1-P2, P2-P3, P3-P4, P4-P5), enquanto o método Permutação Espaço-Tempo identificou apenas um período (P2-P3). O modelo Permutação Espaço-Tempo identificou conglomerados muito maiores que o modelo Poisson no período P5 (região noroeste), enquanto o modelo Poisson identificou um conglomerado maior no período P1-P2 (região sudoeste).

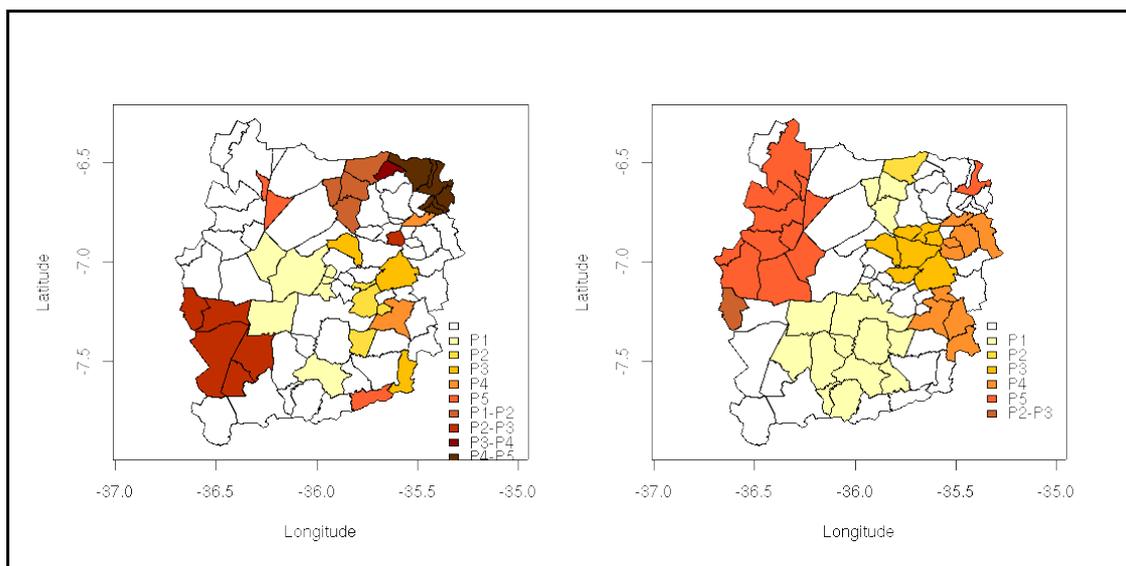


Figura 3. Resultados do método Scan espaço-temporal usando os modelos Poisson (esquerda) e Permutação Espaço-Tempo (direita) para a quinta simulação

CONCLUSÃO

Foi avaliado o desempenho do método Scan espaço-temporal utilizando os modelos Poisson e Permutação Espaço-Tempo por meio de simulações para uma região hipotética. Foram realizadas 50 simulações com cinco períodos de tempo cada. Os conglomerados ocorridos em um único período foram melhor identificados pelo modelo Permutação Espaço-Tempo, enquanto os conglomerados persistentes (espaço-temporais) foi melhor identificada pelo modelo Poisson. Assim, foi perceptível a maior sensibilidade do modelo Poisson, que foi mais adequado aos resultados observados com os mapas de risco, os quais indicavam haver persistência de conglomerados em alguns períodos consecutivos.

REFERÊNCIAS

- [1] COSTA, M.A.; ASSUNÇÃO, R.M. Uma análise de desempenho dos métodos Scan e BESAG&NEWELL na detecção de conglomerados espaciais. In: V Simpósio Brasileiro de Geoinformática - GEOINFO, 2003, São José dos Campos - SP. Anais do V Simpósio Brasileiro de Geoinformática, 2003.
- [2] KULLDORFF, M.; NAGARWALLA, N. Spatial disease clusters: detection and inference. *Statistics in Medicine*, v. 14, p. 799-810, 1995.
- [3] MOURA, F.R. *Detecção de Clusters Espaciais Via Algoritmo Scan Multi-objetivo*. 2006. 51 f. Dissertação (Mestrado em Estatística) - Universidade Federal de Minas Gerais, Belo Horizonte.
- [4] COSTA, M.A.; SCHRERRER, L.R.; ASSUNÇÃO, R.M. Detecção de conglomerados espaciais com geometria arbitrária. In: 17o SINAPE (Simpósio Nacional de Probabilidade e Estatística), 2006, Caxambu. Meio óptico, 2006.
- [5] KULLDORFF, M. SaTScan™ User Guide for version 8.0. February, 2009. Disponível em: <<http://www.satscan.org/>>. Acesso em: 01 abril 2009.
- [6] KULLDORFF, M. A spatial Scan statistic. *Communications in Statistics: Theory and Methods*, v. 26, p. 1481-1496, 1997.
- [7] KULLDORFF, M.; NAGARWALLA, N. Spatial disease clusters: detection and inference. *Statistics in Medicine*, v. 14, p. 799-810, 1995.
- [8] SILVA, A. H. A. ; MARINHO, P. R. D. ; LUCENA, S. E. F. ; MORAES, R. M. . Utilização de métodos estatísticos na análise da distribuição espacial da incidência de dengue no município de João Pessoa no período de 2001 a 2006. In: VII Encontro de Matemática Aplicada e Computacional (ERMAC'2007), 2007, Recife. Anais do VII Encontro de Matemática Aplicada e Computacional (ERMAC'2007), 2007.
- [9] LUCENA, S.E.F.; MORAES, R.,M. Análise do desempenho dos métodos Scan e Besag e Newell para identificação de conglomerados espaciais do dengue no município de João Pessoa entre os meses de janeiro de 2004 e dezembro de 2005. *Boletim de*

Ciências Geodésicas, v. 15, p. 544-561, 2009.

[10] COULSTON, J.W.; RIITERS, K.H. Geographic analysis of forest health indicators using spatial Scan statistics. *Environmental Management*, v. 31, n. 6, p. 764-773, 2003.

[11] KULLDORFF, M. Prospective time-periodic geographical disease surveillance using a Scan statistic. *Journal of the Royal Statistical Society, A* 164, p. 61-72, 2001.

[12] R DEVELOPMENT CORE TEAM. R: A Language and Environment for Statistical Computing. Vienna, Austria, 2007. [online] Disponível em: <http://www.R-project.org>. Acesso em: 05 fevereiro 2007.