



Fundação Oswaldo Cruz
Escola Nacional de Saúde Pública
Departamento de Epidemiologia

Estatística Espacial

Marilia Sá Carvalho

Construção do SIG



- Conceitos gerais
 - Conceitos gerais
 - Aplicações em saúde
 - Conceitos estatísticos fundamentais
 - Tipos de dados

- Áreas
 - Interpolação em áreas
 - Flutuação de pequenas áreas
 - Testes de aglomerados
 - Autocorrelação
 - Indicadores Locais
 - Regressão
 - Espaço-tempo

O que é estatística espacial

- Análise de dados espaciais: em oposição a análise de dados em geral, focaliza-se as técnicas onde se considera explicitamente a localização espacial
- Defini-se “*análise estatística espacial quando os dados são espacialmente localizados e se considera explicitamente a possível importância de seu arranjo espacial na análise ou interpretação dos resultados*” (Bailey & Gatrell, 1995).

Dependência espacial

- Quais as possíveis implicações de não considerar a localização espacial na modelagem?
- *“todas as coisas são parecidas, mas coisas mais próximas se parecem mais que coisas mais distantes”* (Tobler, 1979)

Dependência espacial

- *“Independência é um pressuposto muito conveniente que faz grande parte da teoria estatística matemática tratável. Entretanto, modelos que envolvem dependência estatística são freqüentemente mais realísticos. [...] dados espaciais, onde a dependência está presente em todas as direções e fica mais fraca a medida em que aumenta a dispersão na localização dos dados.” (Cressie,1991)*

Dependência espacial

- Generalizando: a maior parte das ocorrências, sejam estas naturais ou sociais, apresentam entre si uma **relação** - semelhança ou inibição - que depende da **distância**.

Desenho do Estudo

- ❑ Sensibilidade, especificidade e precisão
- ❑ Proporção entre medidas em mapa e medidas reais

Características	Escala ↗	
Resolução	capacidade de distinguir pontos adjacentes	↗
Homogeneidade	característica da distribuição estatística	↗
Instabilidade	presença de flutuação aleatória	↗
Dado	disponibilidade	↘

Aplicações em saúde

- Mapeamento de doenças: descrição do processo espacial de distribuição das doenças, visando vigilância, predição de epidemias, etc.
- Estudos ecológicos: estudar a relação entre incidência de doenças e potenciais fatores etiológicos, seja no campo da análise exploratória visando definir hipóteses (formulação clássica), ou apontar medidas preventivas.

Aplicações em saúde

- Cluster: identificação de focos de doença ou avaliação de aumento de risco ao redor de fonte suspeita de risco ambiental; detecção de epidemia.
- Avaliação e monitoramento ambiental: estimativa e monitoramento da distribuição espacial de fatores ambientais relevantes para a saúde.
- Uso de serviços de saúde: acesso, mercados, qualidade da assistência

Mapeamento de doenças

- O objetivo geral é avaliar a variação geográfica na ocorrência das doenças visando identificar diferenciais de risco, orientar a alocação de recursos, levantar hipóteses etiológicas.

Mapeamento de doenças

- Os métodos tem como objetivo produzir um mapa “limpo”, sem o “ruído” gerado pela flutuação aleatória dos pequenos números, e controlando as diferenças na estrutura demográfica.
- São usualmente aplicados aos dados resultantes de contagens de casos em áreas administrativas ou de indicadores (taxas e proporções).
- Também são aplicados a dados pontuais, usualmente trabalhados sob forma de superfícies de risco, ou de risco relativo.

Estudos ecológicos

- Essencialmente modelos de regressão, onde se busca explicar a variação na incidência da doença através de outras variáveis.
- O modelo se complica pela necessidade de controlar simultaneamente o processo espacial, variáveis explicativas e variáveis de confusão.

Estudos ecológicos

- Classicamente aplica-se a dados agregados em áreas.
- Pode-se entretanto considerar também dados pontuais e misturas de diferentes tipos de dados.

Aglomerado - Cluster

- “*Cluster*”: qualquer agregado de eventos.
- Cluster em estatística multivariada é um resultado de classificação onde se busca definir um grupamento de “semelhantes”.

Aglomerado - Cluster

- Cluster espacial é um agregado de eventos no espaço ou a ocorrência de “taxas semelhantes” em área próximas.
- O objetivo da detecção de cluster espacial é estabelecer a significância de um sobre-risco em um determinado espaço ou tempo e espaço.

Aglomerado

- Estes *aglomerados* podem ser causados por diferentes fatores: agentes infecciosos, contaminação ambiental localizada, efeitos colaterais de tratamentos, etc.

Aglomerado

- Os estudos podem estar dirigidos a buscar evidência de tendência geral à clusterização, ou a um determinado e predefinido agregado.
- Podem ser usados para pontos ou áreas.
- É indispensável controlar para fatores como a distribuição populacional e outras covariáveis que podem criar agregados.

Monitoramento ambiental

- Acompanhamento de potenciais fontes ambientais de problemas de saúde: poluentes químicos, insolação (Raios UV), vegetação, clima, etc.
- Os modelos estatísticos tem por objetivos fazer a predição espacial ou espaço-temporal. Estes processos em geral tem forte correlação espacial e temporal

Monitoramento ambiental

- O interesse pode estar voltado para predição de valores extremos.
- A quantidade e disponibilidade de dados nesta área vem crescendo, com ênfase particular para as imagens de satélite, com resolução e acessibilidade cada vez maior.

Serviços de saúde

- Avaliação de acesso:
 - trajetórias
 - distâncias
 - regionalização

Serviços de saúde

- Áreas de cobertura:
 - Mercados
 - Regionalização
- Qualidade dos serviços:
 - Aspectos demográficos e socioeconômicos.

Conceitos estatísticos fundamentais

- Autocorrelação espacial - caso particular de uma estatística de produtos cruzados do tipo:

$$\Gamma(d) = \sum_{i=1}^n \sum_{j=1}^n w_{ij}(d) \xi_{ij}$$

- Comparando valores obtidos com esperados no caso de independência espacial obtém-se (ou não) evidências da estrutura de dependência.

Conceitos estatísticos fundamentais

- Caso haja evidência de estrutura espacial, o postulado de independência das amostras, base da maior parte dos procedimentos de inferência estatística, é inválido.
- Nestes casos os modelos inferenciais devem levar em conta explicitamente o espaço em suas formulações.
- Dados espaciais não são um conjunto de amostras independentes.

Conceitos estatísticos fundamentais

- Podem ser analisados como uma única realização de um processo estocástico. A diferença em relação a amostragem tradicional, em que cada observação traz uma informação independente, é que todas as observações são utilizadas de forma conjunta para descrever o padrão do fenômeno estudado.

Conceitos estatísticos fundamentais

- Para cada ponto u de uma região contínua em \mathfrak{R}^2 , os valores inferidos de um atributo z – $\hat{z}(u)$ – são realizações de um processo $\{Z(u), u \in A\}$.

Conceitos estatísticos fundamentais

- Estacionariedade
 - as propriedades estatísticas da variável independem de sua localização absoluta, ou seja, a média e a variância são constantes em qualquer sub-área e a covariância entre dois pontos quaisquer depende somente de sua localização relativa.

Conceitos estatísticos fundamentais

- Isotropia
 - se, além de estacionário, a covariância depende somente da distância entre os pontos e não da direção entre eles.

Tipos de dados

- Dados de casos (eventos) - processos pontuais: usualmente se dispõe da localização pontual (coordenadas) da residência de casos de doença ou de controles da população de risco. Covariáveis do indivíduo podem ser medidas.

Tipos de dados

- Áreas - pode-se subdividir em dois sub-grupos:
 - Áreas irregulares - em geral contagens de casos ou populações em divisões administrativas, indicadores socioeconômicos
 - Áreas regulares - medidas em grade regular, como nas imagens de satélite

Tipos de dados

- Dados de amostras pontuais - geoestatística: medidas, em geral de natureza ambiental, tomadas em locais amostrados.
- Referência:

Diggle, P.J. & Ribeiro, P.J. (2000) Model-based Geostatistics. ABE, 14° SINAPE.

Disponível em

<http://www.est.ufpr.br/~paulojus/sinape/>

Tipos de dados

- Misturas de diferentes tipos estão presentes em um mesmo estudo.
- Alguns métodos somente são aplicáveis a um tipo de dado, outros a mais de um.

Tipos de dados

- Em algumas situações pode-se converter o dado de um para outro tipo:
 - pontos ↙ áreas
 - áreas ↙ amostra

Análise exploratória

- Descrição dos dados de forma a contribuir para o desenvolvimento de hipóteses e modelos.
- Caracterizam-se por poucos pressupostos a priori e são resistentes a valores aberrantes (técnicas robustas).

Análise exploratória

- Envolverem, além da visualização, alguma manipulação dos dados, sendo difícil estabelecer a fronteira entre visualização, análise exploratória e modelagem.
- Gráficos dinâmicos - selecionar no mapa e analisar estatística, identificar valores estranhos no mapa.
- Ex: seleção de sub-regiões, análise de vizinhança.

Modelagem

- Nenhum modelo é correto.
- A realidade só pode ser apreendida através de simplificações, denominadas modelos.
- Modelar é a arte de buscar entender e explicar o **essencial** do fenômeno em estudo.
- Modelos são a ponte entre a realidade e a teoria.
- **Modela-se processos não dados!**