March 19, 2002

University of Wisconsin – Madison
Department of Biostatistics and Medical Informatics

Ronald E. Canganon and Murray K. Clayton

Clustered Disease Rates
A Hierarchical Model for Spatially
Hierarchical Model for Spatial Clustering

http://www.biostat.wisc.edu/~ronald
- Inter-neighbors' area, perimeter.

- Geographic centroid of cell from Census Bureau.

- Population from 1980 U.S. Census.

1978-1982,

- Leukemia cases from State Cancer Registry.

- Census tracts or census blocks.

- Eight counties in Upstate New York.

New York Leukemia Data

Hierarchical Model for Spatial Clustering

Cangnon
Questions to Answer

Hierarchical Model for Spatial Clustering

If so, what are the risks associated with the clusters?

If so, where are the clusters located?

Other definitions possible.

Small areas with high (or low) leukemia rates.

Is there evidence of clustering?

Is there evidence of a nonconstant leukemia rate?

What is the leukemia rate for a cell?
Estimate $d^*$ (disease mapping)

Test $H_0: \sum d = \sum d^*$ (clustering)

Goals of analysis:

- Poisson $d_i \sim \frac{\lambda_i}{n_i}$
- $\frac{\lambda_i}{n_i} = (\frac{n_i}{\lambda_i}) E = \sum d_i$
- $n_i = \text{population at risk in cell } i$
- $d_i = \text{number of cases in cell } i$
- $i = 1, 2, \ldots, \sum N$ cells.

Basic Statistical Model

Hierarchical Model for Spatial Clustering

Waller et al. (1992), Lawson (1993)

Stone (1988)

Model-based statistics

Tangeo (2000)

Excess rate statistics

Bonetti (2000)

Whitelemore et al. (1987)

Distance-based statistics

Tests for Spatial Clustering

Hierarchical Model for Spatial Clustering

Ganagnon
Cangnon and Clayton (2001)
Changnon & Clayton (2000)

Dennison & Holmes (2001)

Knorr-Held & Rafter (2000)

Partitioning methods

Besag, York & Molie (1991); Waller et al. (1997)

Local shrinkage models

Clayton & Kaldor (1987)

Global shrinkage models

Approaches

Bayes or Empirical Bayes

Hierarchical Model for Spatial Clustering

Changnon

Point process models
Cluster 

Cluster 

Cluster 

Cluster 

Cluster 

Cluster 

Cluster 

Cluster 

Cluster
\( E_1, \ldots, E_N \sim \text{iid } N(0, 1) \)

Extra-Poisson variation: \( \phi \)

\[ \theta \sum_{k} I_{\{h \in c \}} \]

Note: \( \{h \in c \} \) is the parameter of interest.
Hierarchical Model for Spatial Clustering
Discrete uniform prior for \( \theta \).

Finite number of potential clusters available.

\( \mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_r \) i.i.d. \( p(c) \).

\[ p(0.25) = 0.99, \]

We use \( \theta \) must be fixed.

\( (\mathcal{C}_1, \mathcal{C}_2, \ldots, \mathcal{C}_r, \theta) \) i.i.d. \( \mathcal{N}(0, \sigma^2) \).

Prior Distribution for \( \mathcal{C} \). Hierarchical Model for Spatial Clustering

Gangnan
empty circles.

Flat prior on center/radius of circle adjusted to avoid

Cell belongs to cluster if centroid falls inside circle.

Maximum Geographical Radius.

Radius ranging from 0 km up to 20 km.

Circular clusters centered at cell centroids.

Data

Potential Clusters for New York

Hierarchical Model for Spatial Clustering

Gangnon
Approximate Gibbs sampler serves as proposal

Approximate for likelihood.

Normal priors conjugate to normal

- Parameters $\lambda$, $\theta_1$, $\gamma_1$, $\epsilon_1$, $\epsilon_2$, $\epsilon_3$, $\epsilon_4$,

(1995)

General approach to inference given in Gelman et al.

Hierarchical Poisson Generalized linear model.

Suppose $\gamma_1$, $\gamma_2$, $\gamma_3$ are known.
update.

Conjugate Gamma prior leads to (exact) Gibbs

Parameter $\tau$. 

Hierarchical Model for Spatial Clustering

Gonzalez
Select one of the available transitions at random.

CHANGE the composition of a cluster.

DROP a cluster.

ADD a new cluster.

Possible transitions between models:

(Casson, 1997) to transition between models.

Use reversible jump MCMC (RJMCMC) algorithm.

In reality, \( y \) and \( c_1, \ldots, c_N \) are unknown.

Posterior Simulation

Hierarchical Model for Spatial Clustering
ADD Cluster
DROP Cluster
CHANGE CLUSTER
New York Leukemia Data
Posterior Distribution for $k$
to a Cluster for $k = 3$

Posterior Probability Cell Belongs

Hierarchical Model for Spatial Clustering
Effect for $\beta = 3$

Posterior Mean for Clustering

Hierarchical Model for Spatio-temporal Clustering

Canagahn
is straightforward.

Inclusion of covariate effects and/or temporal effects

of clusters.

Formal methods available for identifying the number

Reversible jump MCMC algorithm for inference.

of model.

Natural prior specification for clustering component

includes extra-Poisson variation.

Proposed spatial model for clustering effects, which

Concluding Remarks

Hierarchical Model for Spatial Clustering

Gangnon