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

Murray K. Clayton

and

Ronald E. Cangaron

August 9, 2001

University of Wisconsin – Madison
Department of Statistics
Department of Biostatistics and Medical Informatics

https://www.biostat.wisc.edu/~ronald
Smoothing may mask clusters.

Smoothed estimators are generally preferred.

- Raw rates tend to be unstable.

- Identification of clusters.

- Examining spatial patterns of disease.

- Useful for:
  - Display disease rate or SMRs as shading on map.

Mapping of Disease Rates
New York Leukemia Data

Raw Rates
Lawson (1995), Lawson & Clark (1999),

Point process models •

Dempston & Holmes (2001),


Partitioning methods •

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

Spatial autoregressive models •

Clayton & Kaldor (1987),

Global shrinkage models •

Bayes or Empirical Bayes Approaches

Hierarchical Models for Spatial Clustering

Clayton & Gampe

August 2001
New York Leukemia Data

Shrinkage Estimates for Rates
Proper, but weak $\text{Gamma}$ priors for $\gamma$ and $\tau$.

and $\phi$ and $\psi$ not separately identified by the likelihood.

\[ (1_{\infty} / Y, Y) \sim (N, \epsilon) = \epsilon \]

Spatial heterogeneity effects: $\epsilon$

\[ (\mathbf{d} / Y, Y) \sim (N \phi, \cdots, \phi) = \phi \]

Spatial autocorrelation effects: $\phi$

A flat prior for $\eta$ (or $g$).

Fixed effect: $\rho$ (could replace with $x$).

\[ \log \psi + \psi + \eta = (d) \]

\[ \text{Poisson}(\mu) \sim \mathcal{O} \]

*Spatial Model of Waller et al. (1997)*

Hierarchical Models for Spatial Clustering

Ganapathy G. Chilamkurti
Proposed Model for Clustering
Can serve as an importance sample for other priors.

- Discrete uniform prior for $k$.

- Finite number of potential clusters available.

\[
c_1, c_2, \ldots, c_k \text{ i.i.d. } p(c) \cdot
\]

We use $\theta_0 = 0.355$ so that $P(0.25 \leq \theta^a \leq 0.99) = 0.15$.

We must be fixed.

\[
\theta_0 \text{ i.i.d. Normal distribution for } \theta_0, \ldots, \theta_1, \ldots, \theta_k \cdot
\]

Prior Distribution for $\theta$. 

Hierarchical Models for Spatial Clustering

Canunnion & Cliellion
• Place prior on center/radius of circle adjusted to avoid empty circles.

• Cell belongs to cluster if centroid falls inside circle.

Maximum Geographic Radii:

• Radii ranging from 0 km up to 20 km.

• Circular clusters centered at cell centroids.

Potential Clusters for New York Data
Conjugate Gamma prior leads to Gibbs update.

- Parameter \( \tau \).

- Distribution for Metropolis algorithm.

Approximate Gibbs sampler serves as proposal

Likelihood.

Normal priors conjugate to normal approximation for

- Parameters \( \mu, \theta_1, \ldots, \theta_k, \epsilon_1, \ldots, \epsilon_n \).

- General approach to inference given in Gelman et al. (1995).

- Hierarchical generalized linear model.

- Suppose \( c_1, \ldots, c_k \) are known.

Posterior Simulation

Hierarchical Models for Spatial Clustering

Gangnan & Chajlin
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:

1997) to transition between models.

Use reversible jump MCMC (RJMCMC) algorithm (Green,

In reality, \( k \) and \( c_i \), \( \ldots \), \( c_h \) are unknown.

Posterior Simulation
New York Leukemia Data

Probability Cell Belongs to a Cluster (k>0)
New York Leukemia Data

Posterior Mean for Cluster-Related Risk (k=3)
New York Leukemia Data

Posterior Means for Rates (k=3)
New York Leukemia Data

Ratio of Posterior Mean Rates ($k=3/k=0$)
Remarks

Concluding Remarks

- Inclusion of covariate effects and/or temporal effects
- Number of clusters
- Formal and informal methods available for identifying the
  Reversible jump MCMC algorithm for inference
- Natural prior specification for clustering component of model
- Extra-Poisson variation
- Proposed spatial model for both clustering effects and

Hierarchical Models for Spatial Clustering

Ganagnon E. Cluytens
\[
\frac{\max_j \frac{V_j}{l_j}}{l_j - l_{j+1}} \cdot \frac{\mathcal{V}}{\mathcal{D}} = \mathcal{V}_m
\]

- Probability of selecting cluster \( i, j \)
- Select cluster radius uniformly from available radii.
- Center the cluster at centroid of that cell.
- Select a cell by throwing a dart at study region.

"Uniform Prior for Clusters"
Propose cluster \( c \) with probability proportional to posterior density.

Propose cluster \( c \) with posterior probability proportional to posterior.

\[
\frac{\theta | c}{\theta |} \cdot \frac{1 + \theta^2 \mathbb{E}}{\theta^2 \mathbb{E} - 2 \theta \mathbb{O}} = \frac{\theta^2}{\theta}
\]

Find posterior mode for the relative risk.

For each potential cluster \( c \),

Need to propose a cluster \( c \) and risk \( 1 + 1 \).

Proposals for the ADD Step
Acceptance probability is easily calculated.

ADD & DROP steps reverse each other.

Equivalent to setting $\theta = 0$.

Propose model without cluster $c_i$.

Assume cluster $k$ is chosen (without loss of generality).

Select one of the $k$ current clusters at random.

Proposals for the DROP Step
Proposed new cluster is always accepted.

Density:

Select cluster \( c \) with probability proportional to posterior.

Model including that cluster with relative risk \( \theta \).

For each potential cluster \( c \), find the posterior density of the

Drop cluster \( c \) from the model.

Keep the current value of \( \theta \) fixed.

Assume cluster \( k \) is chosen (without loss of generality).

Select one of the \( k \) current clusters at random.

Proposals for the CHANCE Step
Expected Cluster Risks

New York Leukemia Data

Hierarchical Models for Spatial Clustering

Gangnon & Clayton