Approaches for Multiple Disease Mapping: MCAR and SANOVA

Sudipto Banerjee

Division of Biostatistics, University of Minnesota

November 3, 2009
Outline

- Introduction
- MCAR: Multivariate Conditionally Auto-Regressive model
- SANOVA: Smoothed ANalysis Of VAriance
- MCAR vs SANOVA: modelling correlations vs means
- Simulation experiments and analysis of Minnesota 3-cancer data
- Discussion and future work
Disease mapping:
- Describe the geographic variation of disease
- Generate hypotheses about the possible causes for differences in risk of disease.
Disease mapping:
- Describe the geographic variation of disease
- Generate hypotheses about the possible causes for differences in risk of disease.

Standard mortality ratio: \( SMR = \frac{Y_i}{E_i} \)
- \( Y_i \) is the observed number of deaths in region \( i \).
- \( E_i \) is the expected number of deaths in region \( i \).

What “factors” are accountable for discrepancies between \( Y_i \)'s and \( E_i \)'s?
Map of raw standard mortality ratios (SMR) of oesophagus cancer between 1991 and 1998 in counties of Minnesota, USA.
Background: Modelling of a single disease

- Let $n$ be the number of counties
- For counts of (rare) disease, poisson regression model:

$$Y_i | \mu_i \sim \text{Poisson} \{E_i \exp(\mu_i)\} \quad i = 1, \ldots, n.$$  

$$\mu_i = \mathbf{x}_i' \beta + \phi_i.$$  

- The $\mathbf{x}_i$’s are explanatory, region-level spatial covariates, having parameter coefficients $\beta$.
- How do we model $\phi_i$’s?
Define:

- $\phi_i$: univariate spatial random variable, $i = 1, \cdots, n$;
- $\phi = (\phi_1, \cdots, \phi_n)'$;
- $i \sim j = \text{region } i \text{ is a neighbor of region } j$;
- $m_i = \text{number of neighbors of region } i$. Let $D = \text{diag}\{m_i\}$
- $W = \text{adjacency matrix of the map}$.  

Brook's Lemma: 

$\phi \sim N(\alpha \sum_{i} \sim j w_{ij} \phi_j, 1)$. 

Aside: $D - \alpha W$ is known as the "Laplacian" of a "graph" in combinatorics (Brualdi and Ryser, 1991).
Define:

- $\phi_i$: univariate spatial random variable, $i = 1, \cdots, n$;
- $\phi = (\phi_1, \cdots, \phi_n)'$;
- $i \sim j = \text{region } i \text{ is a neighbor of region } j$;
- $m_i = \text{number of neighbors of region } i$. Let $D = \text{diag}\{m_i\}$
- $W = \text{adjacency matrix of the map}$.

\[
p(\phi_i \mid \phi_j, j \neq i, \tau) = N \left( \alpha \sum_{i \sim j} \frac{w_{ij}}{w_{i+}} \phi_j, \frac{1}{\tau m_i} \right);
\]

Brook's Lemma: $\Rightarrow \phi \sim N(0, \tau^{-1}(D - \alpha W)^{-1})$. 

Aside: $D - \alpha W$ is known as the "Laplacian" of a "graph" in combinatorics (Brualdi and Ryser, 1991).
Define:

- \( \phi_i \): univariate spatial random variable, \( i = 1, \ldots, n \);
- \( \boldsymbol{\phi} = (\phi_1, \ldots, \phi_n)' \);
- \( i \sim j = \text{region } i \text{ is a neighbor of region } j \);
- \( m_i = \text{number of neighbors of region } i \). Let \( D = \text{diag}\{m_i\} \)
- \( W = \text{adjacency matrix of the map} \).

\[
p(\phi_i | \phi_j, j \neq i, \tau) = N \left( \alpha \sum_{i \sim j} \frac{w_{ij} \phi_j}{w_i}, \frac{1}{\tau m_i} \right);
\]

Brook’s Lemma: \( \Rightarrow \phi \sim N(0, \tau^{-1}(D - \alpha W)^{-1}) \).

Aside: \( D - \alpha W \) is known as the “Laplacian” of a “graph” in combinatorics (Brualdi and Ryser, 1991).
\[ \lambda_{\text{min}} \text{ and } \lambda_{\text{max}} \text{ are the minimum and maximum eigenvalues of } D^{-1/2}WD^{-1/2}. \]

- In fact, \( \lambda_{\text{min}} < 0 \) and \( \lambda_{\text{max}} = 1 \).

- \( \alpha \in (\lambda^{-1}_{\text{min}}, \lambda^{-1}_{\text{max}}) \) ensures proper distribution for \( \phi \) (Cressie, 1993; Sun et al., 2000).

- \( \alpha = 1 \) implies improper (singular) distribution (Besag et al. 1991).

\[
p(\phi_1, \phi_2, \ldots, \phi_n \mid \tau) \propto (\tau)^{(n-G)/2} \exp \left\{ -\frac{\tau}{2} \phi'(D - W)\phi \right\},
\]

where \( G \) is the number of “islands” in the map. In fact, \( n - G \) is the rank of \( D - W \).
Hierarchical Bayesian modeling:

- First stage: likelihood for observation data

- Second stage: prior distributions for fixed effects $\beta$ and random effects $\psi = (\psi_1, \ldots, \psi_N)'$.

- Estimate model using Markov chain Monte Carlo (MCMC) methods
$Y_{ij}$ is the disease count for disease $j$ in county $i$. 

For counts of (rare) diseases, Poisson regression model:

$$Y_{ij} \sim \text{Poisson}\left(E_{ij}e^{x_{ij}'\beta_j + \phi_{ij}}\right), \quad i = 1, \ldots, n, \quad j = 1, \ldots, p.$$ 

How do we model the $\phi_{ij}$’s?
• $Y_{ij}$ is the disease count for disease $j$ in county $i$.

• $x_{ij}$ are explanatory, region-level spatial covariates for disease $j$ having parameter coefficients $\beta_j$. 
\( Y_{ij} \) is the disease count for disease \( j \) in county \( i \).

\( x_{ij} \) are explanatory, region-level spatial covariates for disease \( j \) having parameter coefficients \( \beta_j \).

For counts of (rare) diseases, poisson regression model:

\[
Y_{ij} \overset{ind}{\sim} \text{Poisson}(E_{ij}e^{x_{ij}'\beta_j + \phi_{ij}}), \quad i = 1, \ldots, n, \ j = 1, \ldots, p.
\]
$Y_{ij}$ is the disease count for disease $j$ in county $i$.

$x_{ij}$ are explanatory, region-level spatial covariates for disease $j$ having parameter coefficients $\beta_j$.

For counts of (rare) diseases, poisson regression model:

Multiple disease model

$$Y_{ij} \overset{ind}{\sim} Poisson(E_{ij}e^{x'_{ij}\beta_j + \phi_{ij}}), \quad i = 1, \ldots, n, \quad j = 1, \ldots, p.$$ 

How do we model the $\phi_{ij}$'s?
Associations in multiple disease data:
Associations in multiple disease data:

- Spatial association across regions – now for each disease
- Dependence among multiple diseases (within the same region)
- Cross-spatial associations among multiple diseases in different regions
Define $\phi_j = (\phi_{1j}, \cdots, \phi_{nj})'$; 

$\phi' = (\phi'_1, \cdots, \phi'_p)$.

$$\phi \sim N(0, \Lambda^{-1} \otimes (D - \alpha W)^{-1})$$

Covariance matrix ($p = 2$):

$$
\begin{pmatrix}
\phi_1 \\
\phi_2
\end{pmatrix} \sim N
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{pmatrix}
(D - \alpha W)\Lambda_{11} & (D - \alpha W)\Lambda_{12} \\
(D - \alpha W)\Lambda_{12} & (D - \alpha W)\Lambda_{22}
\end{pmatrix}^{-1}
$$

Note: Common spatial smoothing parameter is used for all diseases.
MCAR model: \( MCAR(\alpha_1, \ldots, \alpha_p, \Lambda) \) (Gelfand and Vounatsou, 2003)

Now every disease has its own smoothing parameter.

Covariance matrix \((p = 2)\).

\[
\begin{pmatrix}
\phi_1 \\
\phi_2
\end{pmatrix}
\sim \mathcal{N}\left(
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{pmatrix}
(D - \alpha_1 W)\Lambda_{11} & * \\
* & (D - \alpha_2 W)\Lambda_{22}
\end{pmatrix}^{-1}
\right),
\]

No additional parameter to control smoothness of cross-correlations.
Gelfand and Vounatsou (2003) recognize the difficulties in specifying covariance matrices of the following form (e.g. with $p = 2$):

\[
\begin{pmatrix}
\phi_1 \\
\phi_2
\end{pmatrix}
\sim
N
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{pmatrix}
(D - \alpha_1 W)\Lambda_{11} & (D - \alpha_3 W)\Lambda_{12} \\
(D - \alpha_3 W)\Lambda_{12} & (D - \alpha_2 W)\Lambda_{22}
\end{pmatrix}^{-1}
\]
Gelfand and Vounatsou (2003) recognize the difficulties in specifying covariance matrices of the following form (e.g. with \( p = 2 \)):

\[
\begin{pmatrix}
\phi_1 \\
\phi_2
\end{pmatrix} \sim N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix}
(D - \alpha_1 W)\Lambda_{11} & (D - \alpha_3 W)\Lambda_{12} \\
(D - \alpha_3 W)\Lambda_{12} & (D - \alpha_2 W)\Lambda_{22}
\end{pmatrix}^{-1}\right)
\]

Kim, Sun and Tsutakawa (2001): Two-fold CAR model that allows smoothness parameters for cross-correlations for two diseases. Also work by Sain and Cressie (2002). Extremely difficult to generalize to \( p > 2 \).
Gelfand and Vounatsou (2003) recognize the difficulties in specifying covariance matrices of the following form (e.g. with \( p = 2 \)):

\[
\begin{pmatrix}
\phi_1 \\
\phi_2
\end{pmatrix}
\sim
N
\left(
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\begin{pmatrix}
(D - \alpha_1 W)\Lambda_{11} & (D - \alpha_3 W)\Lambda_{12} \\
(D - \alpha_3 W)\Lambda_{12} & (D - \alpha_2 W)\Lambda_{22}
\end{pmatrix}
\right)^{-1}
\]

Kim, Sun and Tsutakawa (2001): Two-fold CAR model that allows smoothness parameters for cross-correlations for two diseases. Also work by Sain and Cressie (2002). Extremely difficult to generalize to \( p > 2 \).

Jin et al. (2005): Conditional approach to building 2-CAR models: \( [\phi_1][\phi_2 | \phi_1] \).
MCAR models

Sanova vs MCAR
Borrow ideas from factor analysis and multivariate geostatistics?
Borrow ideas from factor analysis and multivariate geostatistics?

For $i = 1, \ldots, n$ and $j = 1, \ldots, p$:

$$\phi_{ij} = a_{j1}u_{i1} + a_{j2}u_{i2} + \cdots + a_{jp}u_{ip}$$
Borrow ideas from factor analysis and multivariate geostatistics?

For \( i = 1, \ldots, n \) and \( j = 1, \ldots, p \):

\[
\phi_{ij} = a_{j1}u_{i1} + a_{j2}u_{i2} + \cdots + a_{jp}u_{ip}
\]

Letting \( \phi'_j = (\phi_{1j}, \ldots, \phi_{nj}) \),

\[
\phi_j = a_{j1}u_1 + a_{j2}u_2 + \cdots + a_{jp}u_p
\]
Borrow ideas from factor analysis and multivariate geostatistics?

For $i = 1, \ldots, n$ and $j = 1, \ldots, p$:

$$\phi_{ij} = a_{j1}u_{i1} + a_{j2}u_{i2} + \cdots + a_{jp}u_{ip}$$

Letting $\phi'_j = (\phi_{1j}, \ldots, \phi_{nj})$,

$$\phi_j = a_{j1}u_1 + a_{j2}u_2 + \cdots + a_{jp}u_p$$

$u_j$’s are “latent” variables (no dimension reduction)

$a_{jk}$’s are unknown coefficients (weights) – not varying over regions (but could)
Case 1: \( p \) independent and identical latent CAR effects

- Assume \( p \) latent independent CAR effects \( u_j \) (one for each disease):

\[
  u_j \sim N_n \left( 0, (D - \alpha W)^{-1} \right), \quad j = 1, \ldots, p.
\]
Case 1: \( p \) independent and identical latent CAR effects

- Assume \( p \) latent independent CAR effects \( u_j \) (one for each disease):

\[
    u_j \sim N_n \left( 0, (D - \alpha W)^{-1} \right), \ j = 1, \ldots, p.
\]

- Let \( \phi' = (\phi'_1, \ldots, \phi'_p) \), \( A = \{a_{jk}\} \) and \( \Sigma = AA' \).
Case 1: \( p \) independent and identical latent CAR effects

- Assume \( p \) latent independent CAR effects \( u_j \) (one for each disease):

\[
    u_j \sim N_n \left( \mathbf{0}, \left( D - \alpha W \right)^{-1} \right), \ j = 1, \ldots, p.
\]

- Let \( \phi' = (\phi'_1, \ldots, \phi'_p) \), \( A = \{a_{jk}\} \) and \( \Sigma = AA' \).

- The joint distribution of \( \phi \) is

\[
    \phi \sim N \left( \mathbf{0}, \Sigma \otimes \left( D - \alpha W \right)^{-1} \right).
\]
Case 1: \( p \) independent and identical latent CAR effects

- Assume \( p \) latent independent CAR effects \( u_j \) (one for each disease):

\[
    u_j \sim N_n \left( 0, (D - \alpha W)^{-1} \right), \quad j = 1, \ldots, p.
\]

- Let \( \phi' = (\phi'_1, \ldots, \phi'_p) \), \( A = \{a_{jk}\} \) and \( \Sigma = AA' \).

- The joint distribution of \( \phi \) is

\[
    \phi \sim N \left( 0, \Sigma \otimes (D - \alpha W)^{-1} \right).
\]

- This is the separable MCAR distribution \( MCAR(\alpha, \Sigma) \) distribution.
Case 2: $p$ independent but not identical latent CAR effects

Assume that the latent CAR effects $u_j$ are independent and not identical.

$$u_j \sim N \left( 0, \left( D - \alpha_j W \right)^{-1} \right), \quad j = 1, \ldots, p.$$
Case 2: \( p \) independent but not identical latent CAR effects

- Assume that the latent CAR effects \( u_j \) are independent and not identical.

\[
  u_j \sim N \left( 0, (D - \alpha_j W)^{-1} \right), \ j = 1, \ldots, p.
\]

- The joint distribution of \( \phi \) is

\[
  \phi \sim N \left( 0, (A \otimes I_{n \times n}) \Gamma^{-1} (A' \otimes I_{n \times n}) \right),
\]
Case 2: \( p \) independent but not identical latent CAR effects

- Assume that the latent CAR effects \( u_j \) are independent and not identical.

\[
 u_j \sim N \left( 0, \left( D - \alpha_j W \right)^{-1} \right), \ j = 1, \ldots, p.
\]

- The joint distribution of \( \phi \) is

\[
 \phi \sim N \left( 0, \left( A \otimes I_{n \times n} \right) \Gamma^{-1} \left( A' \otimes I_{n \times n} \right) \right),
\]

- \( \Gamma \) is block diagonal with diagonal entries \( \Gamma_j = D - \alpha_j W \), \( j = 1, \ldots, p \).
Case 2: \( p \) independent but not identical latent CAR effects

- Assume that the latent CAR effects \( u_j \) are independent and not identical.

\[
  u_j \sim N \left(0, (D - \alpha_j W)^{-1}\right), \ j = 1, \ldots, p.
\]

- The joint distribution of \( \phi \) is

\[
  \phi \sim N \left(0, (A \otimes I_{n \times n}) \Gamma^{-1} (A' \otimes I_{n \times n}) \right),
\]

- \( \Gamma \) is block diagonal with diagonal entries \( \Gamma_j = D - \alpha_j W \), \( j = 1, \ldots, p \).

- This is the non-separable MCAR, \( MCAR(\alpha_1, \ldots, \alpha_p, \Sigma) \) distribution.
Case 3: \( p \) dependent and not identical latent CAR effects

Assume that the latent CAR effects \( u' = (u'_1, \ldots, u'_p) \) are such that:

\[
u \sim N \left( 0, (I_{p \times p} \otimes D - B \otimes W)^{-1} \right),
\]

where \( B = \{b_{jk}\} \) is a \( p \times p \) symmetric matrix.
Case 3: \( p \) dependent and not identical latent CAR effects

Assume that the latent CAR effects \( u' = (u'_1, \ldots, u'_p) \) are such that:

\[
    u \sim N \left( 0, \left( I_{p \times p} \otimes D - B \otimes W \right)^{-1} \right),
\]

where \( B = \{b_{jk}\} \) is a \( p \times p \) symmetric matrix.

The joint distribution of \( \phi \) is

\[
    \phi \sim N \left( 0, \left( A \otimes I_{n \times n} \right) \left( I_{p \times p} \otimes D - B \otimes W \right)^{-1} \left( A \otimes I_{n \times n} \right)' \right).
\]
**Case 3:** \( p \) dependent and not identical latent CAR effects

- Assume that the latent CAR effects \( u' = (u'_1, \ldots, u'_p) \) are such that:

\[
u \sim N \left(0, (I_{p \times p} \otimes D - B \otimes W)^{-1}\right),\]

where \( B = \{b_{jk}\} \) is a \( p \times p \) symmetric matrix.

- The joint distribution of \( \phi \) is

\[
\phi \sim N \left(0, (A \otimes I_{n \times n}) (I_{p \times p} \otimes D - B \otimes W)^{-1} (A \otimes I_{n \times n})'\right).
\]

- We call this the \( MCAR(B, \Sigma) \) distribution
Method behind this “madness”? 

- The $MCAR(B, \Sigma)$ distribution can be expressed (e.g. $p = 2$ diseases) as

$$\phi \sim N \left(0, \begin{pmatrix} (D - \alpha_{11} W)\Lambda_{11} & (D - \alpha_{12} W)\Lambda_{12} \\ (D - \alpha_{12} W)\Lambda_{12} & (D - \alpha_{22} W)\Lambda_{22} \end{pmatrix}^{-1} \right).$$

- Result holds for general $p$, i.e. matrices of the form
  \{(D - \alpha_{ij} W)\Lambda_{ij}\}

- There exists a bijection:

$$\left(\{a_{jk}\}, \{b_{jk}\}\right) \Leftrightarrow \left(\{\Lambda_{ij}\}, \{\alpha_{ij}\}\right)$$
Computation Notes

- **Restriction on** $B$:
  - $\xi_i = \text{eigenvalues of } D^{-1/2}WD^{-1/2}, \ i = 1, \ldots, n; (\xi_{\text{min}} < 0)$
  - $\zeta_j = \text{eigenvalues of } B, \ j = 1, \cdots, p$
  - $\xi_i \times \zeta_j = \text{eigenvalues for } B \otimes (D^{-1/2}WD^{1/2})$
  - $I_p \otimes D - B \otimes W$ is p.d. iff $\zeta_j \in (1/\xi_{\text{min}}, 1)$.
Computation Notes

- **Restriction on** $B$:
  - $\xi_i = \text{eigenvalues of } D^{-1/2} W D^{-1/2}, \ i = 1, \ldots, n; (\xi_{\min} < 0)$
  - $\zeta_j = \text{eigenvalues of } B, \ j = 1, \ldots, p$
  - $\xi_i \times \zeta_j = \text{eigenvalues for } B \otimes (D^{-1/2} W D^{1/2})$
  - $I_p \otimes D - B \otimes W$ is p.d. iff $\zeta_j \in (1/\xi_{\min}, 1)$.

- **Update** $B$:
  - Spectral decomposition of $B$: $B = P \Delta P'$;
  - $P = \prod_{i=1}^{p-1} \prod_{j=i+1}^{p} G_{ij}(\theta_{ij})$ where $G_{ij}$ is $I_p$ with $i$-th and $j$-th diagonal elements replaced by $\cos(\theta_{ij})$ and $(i, j)$-th and $(j, i)$-th entry replaced by $\pm \sin(\theta_{ij})$;
  - priors: $\theta_{ij} \sim U(-\pi/2, \pi/2), \ \zeta_j \sim U(1/\xi_{\min}, 1)$
Minnesota Cancer Surveillance data used;

Counties in Minnesota are treated as locations: $j = 1, \ldots, 87$.

3 types of diseases considered: cancers of the lung, larynx and oesophagus: $j = 1, 2, 3$.

$E_{ij}$ computed as the expected age adjusted number of deaths due to cancer $j$ in county $i$. 
### Table 4. Posterior summaries of parameters in the MCAR($B$, $\Sigma$) model for the Minnesota cancer data

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Results for the following cancers:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lung</td>
</tr>
<tr>
<td></td>
<td>Median (2.5%, 97.5%)</td>
</tr>
<tr>
<td>$\beta_1, \beta_2, \beta_3$</td>
<td>$-0.093 (-0.179, -0.006)$</td>
</tr>
<tr>
<td>$\Sigma_{11}, \Sigma_{22}, \Sigma_{33}$</td>
<td>$0.048 (0.030, 0.073)$</td>
</tr>
<tr>
<td>$\rho_{12}, \rho_{13}$</td>
<td>$0.277 (-0.112, 0.643)$</td>
</tr>
<tr>
<td>$\rho_{23}$</td>
<td>$0.442 (-0.302, 0.921)$</td>
</tr>
<tr>
<td>$b_{11}, b_{22}, b_{33}$</td>
<td>$0.036 (-0.830, 0.857)$</td>
</tr>
<tr>
<td>$b_{12}, b_{13}$</td>
<td>$0.323 (-0.156, 0.842)$</td>
</tr>
<tr>
<td>$b_{23}$</td>
<td>$0.006 (-0.519, 0.513)$</td>
</tr>
</tbody>
</table>
Do we really need these complex multivariate structures?
Can the data really identify MCAR’s elaborate dependence structures?
What if we want to work with an improper MCAR?
Can we have a simpler and more interpretable approach?
A new approach to multiple disease mapping

Two-way ANOVA with one observation per cell

<table>
<thead>
<tr>
<th>Cancer 1</th>
<th>County 1</th>
<th>County 2</th>
<th>\cdots</th>
<th>County n</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{Cancer 1}</td>
<td>*</td>
<td>*</td>
<td>\cdots</td>
<td>*</td>
</tr>
<tr>
<td>\text{Cancer 2}</td>
<td>*</td>
<td>*</td>
<td>\cdots</td>
<td>*</td>
</tr>
<tr>
<td>:</td>
<td>:</td>
<td>:</td>
<td>\vdots</td>
<td>:</td>
</tr>
<tr>
<td>\text{Cancer p}</td>
<td>*</td>
<td>*</td>
<td>\cdots</td>
<td>*</td>
</tr>
</tbody>
</table>
### Two-way ANOVA Table

<table>
<thead>
<tr>
<th>Effects</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>$p - 1$</td>
</tr>
<tr>
<td>County</td>
<td>$n - 1$</td>
</tr>
<tr>
<td>Error = Cancer × County</td>
<td>$(n - 1)(p - 1)$</td>
</tr>
<tr>
<td>Total</td>
<td>$np - 1$</td>
</tr>
</tbody>
</table>
Two-way ANOVA table

<table>
<thead>
<tr>
<th>Effects</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>$p - 1$</td>
</tr>
<tr>
<td>County</td>
<td>$n - 1$</td>
</tr>
<tr>
<td>Error=Cancer×County</td>
<td>$(n - 1)(p - 1)$</td>
</tr>
<tr>
<td>Total</td>
<td>$np - 1$</td>
</tr>
</tbody>
</table>

Saturated linear model:

Grand Mean + Cancer Main Effects + County Main Effects + Cancer×County Interaction Effects
A new approach to multiple disease mapping?

- Smoothed ANOVA (Hodges et al, 2007):
  \[ g(E[y]) = A_1 \Theta_1 + A_2 \Theta_2; \]

- \( M_1 \) and \( M_2 \) degrees of freedom for main effects and interactions respectively

- \( A_1 \)'s columns correspond to the grand mean and main effects; \( A_1'A_1 = I_{M_1} \).

- \( A_2 \)'s columns correspond to interactions; \( A_2'A_2 = I_{M_2} \).

- Rather than choosing interaction (and main effects) to include/exclude...Include all these effects, but shrink toward zero.
$A_1$ and $A_2$ are design matrices of contrasts for the unshrunk (overall & main) and shrunk (interaction) effects.

“Smoothing”: $\Theta_2 \sim N(0, \text{diag}(\frac{1}{\eta_j}))$. Usually the interaction effects are shrunk towards zero.

$\Theta_1$ is assigned a flat prior; grand mean and main effects are not smoothed.
- $A_1$ and $A_2$ are design matrices of contrasts for the unshrunk (overall & main) and shrunk (interaction) effects.

- “Smoothing”: $\Theta_2 \sim N(0, diag(\frac{1}{\eta_j}))$. Usually the interaction effects are shrunk towards zero.

- $\Theta_1$ is assigned a flat prior; grand mean and main effects are not smoothed.

- How does the improper CAR fit into SANOVA?
Define $Q = D - W$.

The intrinsic CAR model (Besag et al 1991) with $L_2$ norm has the improper density

$$p(\phi | \tau) \sim \tau^{\frac{n-G}{2}} \exp \left( -\frac{\tau}{2} \phi' Q \phi \right),$$

where $G$ is the number of "islands" (disconnected parts). Assume that $G = 1$ (no major issues for general $G$).
Define $Q = D - W$.

The intrinsic CAR model (Besag et al 1991) with $L_2$ norm has the improper density

$$p(\phi | \tau) \sim \tau^{\frac{n-G}{2}} \exp \left(-\frac{\tau}{2} \phi' Q \phi\right),$$

- $\tau$ is the spatial dispersion parameter.
- $G$ is the number of “islands" (disconnected parts).
- Assume that $G = 1$ (no major issues for general $G$)
\[ Q_1 = 0 \text{ and rank of } Q \text{ is } n - 1. \]
\( Q_1 = 0 \) and rank of \( Q \) is \( n - 1 \).

Let \( Q \) have spectral decomposition \( Q = V \Lambda V' \), \( \Lambda \) is diagonal and \( V \) is orthogonal with columns as eigenvectors.
\[ Q_1 = 0 \] and rank of \( Q \) is \( n - 1 \).

Let \( Q \) have spectral decomposition \( Q = V \Lambda V' \), \( \Lambda \) is diagonal and \( V \) is orthogonal with columns as eigenvectors.

Eigenvector \( \frac{1}{\sqrt{n}} 1_n \) has eigenvalue 0.
• $Q1 = 0$ and rank of $Q$ is $n - 1$.

• Let $Q$ have spectral decomposition $Q = V\Lambda V^\prime$, $\Lambda$ is diagonal and $V$ is orthogonal with columns as eigenvectors.

• Eigenvector $\frac{1}{\sqrt{n}}1_n$ has eigenvalue 0.

• Let $V$’s other columns are orthogonal to $1_n$; call them $V_1$ and their corresponding eigenvalue matrix as $\Lambda_1$. 
$Q1 = 0$ and rank of $Q$ is $n - 1$.

Let $Q$ have spectral decomposition $Q = V \Lambda V'$, $\Lambda$ is diagonal and $V$ is orthogonal with columns as eigenvectors.

Eigenvector $\frac{1}{\sqrt{n}} 1_n$ has eigenvalue $0$.

Let $V$’s other columns are orthogonal to $1_n$; call them $V_1$ and their corresponding eigenvalue matrix as $\Lambda_1$.

$V_1$’s $n - 1$ columns also define contrasts in $\phi$’s: $1' V_1 = 0$. 
Define a new parameter $\Theta = V'\phi$, so

$$p(\Theta | \tau) \sim \tau^{n-G} \exp \left( -\frac{\tau}{2} \Theta' \Lambda \Theta \right),$$
Define a new parameter $\Theta = V' \phi$, so

$$p(\Theta | \tau) \sim \tau \frac{n-G}{2} \exp \left( -\frac{\tau}{2} \Theta' \Lambda \Theta \right),$$

Setting a CAR prior on $\phi \iff$ setting $\Theta$’s prior to $N(0, (\tau \Lambda)^{-1})$. 

SANOVA vs MCAR
• Define a new parameter $\Theta = V' \phi$, so

$$p(\Theta | \tau) \sim \tau^{n-G/2} \exp \left( -\frac{\tau}{2} \Theta' \Lambda \Theta \right),$$

Setting a CAR prior on $\phi \iff$ setting $\Theta$’s prior to $N(0, (\tau \Lambda)^{-1})$.

• Let $\phi' = (\phi_1, \phi_2, \ldots, \phi_n)$, then the SANOVA model

$$\phi = V \Theta$$

$$= \begin{bmatrix} V_1 & \frac{1}{\sqrt{n}} 1_n \end{bmatrix} \begin{bmatrix} \Theta_{Reg} \\ \Theta_{GM} \end{bmatrix}.$$ 

where $\Theta_{REG} \sim N(0, (\tau \Lambda_1)^{-1})$ and the precision $\Lambda_{nn} = 0$ for the overall level, i.e., a flat prior on $\Theta_{GM}$. 
Suppose $\phi$, the spatial effects for $p$ cancers in $n$ counties, arises from a saturated linear regression model:

$$\phi = [\phi_1', \phi_2', \ldots, \phi_n']' = [A_1|A_2] \Theta =$$

$$
\begin{pmatrix}
\frac{1}{\sqrt{np}} 1_{np} & \frac{1}{\sqrt{n}} 1_n \otimes H_{CA} & V_1 \otimes \frac{1}{\sqrt{p}} 1_p & V_1 \otimes H_{CA}^{(1)} \ldots V_1 \otimes H_{CA}^{(p-1)} \\
\end{pmatrix}
\begin{pmatrix}
\Theta_{GM} \\
\Theta_{CA} \\
\Theta_{CO} \\
\Theta_{CO \times CA}
\end{pmatrix}
$$

<table>
<thead>
<tr>
<th>Grand mean</th>
<th>Cancer main effect</th>
<th>County main effect</th>
<th>Cancer x County interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>$np \times 1$</td>
<td>$np \times (p-1)$</td>
<td>$np \times (n-1)$</td>
<td>$np \times (n-1)(p-1)$</td>
</tr>
</tbody>
</table>

Let $\Theta_{CO} \sim N_{n-1}(0, (\tau_0 \Lambda_1)^{-1})$; $\Theta_{CO \times CA}^{(j)} \sim N_{n-1}(0, (\tau_j \Lambda_1)^{-1})$,

where $H_{CA}$ is $p \times (p-1)$ matrix whose columns are contrasts for cancer, $\Lambda_1$ corresponds to $V_1$; $\tau_0, \tau_j > 0$ are unknown and $\tau_0 \Lambda_1, \tau_j \Lambda_1$ are precision matrices.
We can prove that a *priori*,

\[
A_2 \begin{pmatrix}
\Theta_{CO} \\
\Theta_{CO \times CA}^{(1)} \\
\vdots \\
\Theta_{CO \times CA}^{(p-1)}
\end{pmatrix}
\] has precision \( Q \otimes (H_A^{(+) \, \text{diag}(\tau_j)H_A^{(+)\prime}) \),
We can prove that a \textit{priori},

$$A_2 \begin{pmatrix} \Theta CO \\ \Theta CO \times CA^{(1)} \\ \vdots \\ \Theta CO \times CA^{(p-1)} \end{pmatrix} \text{ has precision } Q \otimes (H_A^+ \text{diag}(\tau_j)H_A^+)^\prime,$$

\(A_2\) refers to the columns of the design matrix corresponding to the county main effects and the cancer-by-county interactions.
We can prove that a *priori*,

\[
A_2 \begin{pmatrix}
\Theta_{CO} \\
\Theta_{CO \times CA}^{(1)} \\
\vdots \\
\Theta_{CO \times CA}^{(p-1)} 
\end{pmatrix}
\] has precision \( Q \otimes (H_A^{(+)} \text{diag}(\tau_j) H_A^{(+)\prime}) \),

- \( A_2 \) refers to the columns of the design matrix corresponding to the county main effects and the cancer-by-county interactions
- \( H_A^{(+)} = \left( \frac{1}{\sqrt{p}} 1_p \mid H_{CA} \right) \) is an orthogonal matrix
The marginal precision matrix of \([\phi_1', \phi_2', \ldots, \phi_N']'\) is:

**MCAR**

\[ Q \otimes (V_\Omega D_\Omega V_\Omega') \]

**SANOVA**

\[ Q \otimes (H_A^{(+) \text{diag}(\tau_j) H_A^{(+)})' \]

unknown \quad known \quad unknown
We simulated both normal and Poisson distributed data for different settings of precision parameters. When generating data, we set

\[
H_A^{(+)} = HA_1 = \begin{pmatrix}
1 & -2 & 0 \\
1 & 1 & -1 \\
1 & 1 & 1
\end{pmatrix}
\begin{pmatrix}
\frac{1}{\sqrt{3}} \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
\frac{1}{\sqrt{6}} \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
\frac{1}{\sqrt{2}}
\end{pmatrix}.
\]

We fit three SANova models: SANova with the correct \(H_A^{(+)}\), \(HA_1\); SANova with a somewhat incorrect \(H_A^{(+)}\), \(HA_2\); a variant SANova with a very incorrect \(H_A^{(+)}\), \(HM\), where

\[
HA_2 = \begin{pmatrix}
1 & 1 & 1 \\
1 & -2 & 0 \\
1 & 1 & -1
\end{pmatrix}
\begin{pmatrix}
\frac{1}{\sqrt{3}} \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
\frac{1}{\sqrt{6}} \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
\frac{1}{\sqrt{2}}
\end{pmatrix},
\]

\[
H_M = \begin{pmatrix}
0.56 & -0.64 & -0.52 \\
-0.53 & -0.77 & 0.36 \\
-0.63 & 0.07 & -0.77
\end{pmatrix}.
\]
Six models were compared.

- Three flavours of SANova
  - “SANova Correct” has the $H_A^{(+)}$ used to generate the data.
  - “SANova Incorrect” used a slightly wrong $H_A^{(+)}$ matrix.
  - “SANova Variant” data generated from an MCAR (very wrong $H_A^{(+)}$).

- Compared to an MCAR model with varying values for $\Omega$’s hyper-prior matrix (i.e. the Wishart scale matrix) $R$.

\[
R = I_p \\
R = 0.002 \cdot I_n \\
R = 200 \cdot I_n
\]
Simulation Findings

Metrics used:

- **AMSE**: Average (over \( L = 100 \) replications) MSE of the cell means plugging in the posterior medians

- **MBIAS**: 2.5, 50, & 97.5 percentiles of the average (over replications) bias of the cell means (plugging in the posterior median).

- **PI Rate**: average coverage of the 95% posterior intervals of the cell means.

- **DIC scores**
Findings:

- “SANova Correct” did the best (smallest AMSE & MBIAS, nearly nominal PI Rate).
- For normal data: SANova mostly outperformed MCAR.
- SANova models were comparable but worse than MCAR for good choices of the Wishart’s scale matrix;
- MCAR performed poorly for very bad choices on $R$
- Overall: SANova appears competitive than MCAR and is fairly robust to mis-specifying $H_A^{(+)},$ unlike MCAR which is relatively sensitive to the hyper-prior on $\Omega.$
Maps of raw and fitted SMR's

(a) data: $y_{ij}/E_{ij}$

- (0) $< 0.5$
- (28) 0.5 - 1.0
- (19) 1.0 - 1.2
- (25) 1.2 - 1.5
- (14) 1.5 - 2.0
- (1) $>= 2.0$

(b) fitted: SANOVA with $R_{A1}$

(c) fitted: MCAR with $R_{a1}$ = 1

Lung Cancer

SANOVA vs MCAR
Maps of raw and fitted SMR’s

Esophagus Cancer

(a) data: $y_{ij}/E_{ij}$

- (0) < 0.5
- (28) 0.5 - 1.0
- (19) 1.0 - 1.2
- (25) 1.2 - 1.5
- (14) 1.5 - 2.0
- (1) >= 2.0

(b) fitted: SANova with $H_{A1}$

(c) fitted: MCAR with $R_m = 1$
We extended SANOVA to smooth spatial random effects taking advantage of the spatial structure. For the cases considered, SANOVA does in fact, compete well with MCAR. To improve SANOVA in the future,

- Extend SANOVA to more general MCAR case (ours is intrinsic MCAR).
- SANOVA in spatiotemporal models.
- SANOVA in spatial survival models (e.g., Cox model with spatial effects).
- Further investigations on choice of priors and robustness issues.


Thank You!