WinBUGS Example 1: Lip cancer

Consider the areal data disease mapping model:

\[ Y_i \mid \mu_i \overset{\text{ind}}{\sim} \text{Po}(E_i e^{\mu_i}) , \text{ where} \]
\[ Y_i = \text{observed disease count}, \]
\[ E_i = \text{expected count (known)}, \text{ and} \]
\[ \mu_i = x_i'\beta + \theta_i + \phi_i \]

- \( Y_i \) informs directly only about \( \xi_i \equiv \theta_i + \phi_i \)
- The \( x_i \) are explanatory spatial covariates; typically \( \beta \) has a flat prior.
- The \( \theta_i \) capture heterogeneity among the regions via

\[ \theta_i \overset{iid}{\sim} N(0, 1/\tau_h) , \]
WinBUGS Example 1: Lip cancer

- and the $\phi_i$ capture regional clustering via a conditionally autoregressive (CAR) prior,

\[
\phi_i \mid \phi_j \neq i \sim N(\bar{\phi}_i, 1/(\tau_c m_i)),
\]

where $\bar{\phi}_i = m_i^{-1} \sum_{j \in \partial_i} \phi_j$, $\partial_i$ is the set of “neighbors” of region $i$, and $m_i$ is the number of these neighbors.

- The CAR prior is translation invariant, so typically we insist $\sum_{i=1}^{I} \phi_i = 0$ (imposed numerically after each MCMC iteration).

- Making the reparametrization from $(\theta, \phi)$ to $(\theta, \xi)$, we have the joint posterior

\[
p(\theta, \xi \mid y) \propto L(\xi; y)p(\theta)p(\xi - \theta).
\]
WinBUGS Example 1: Lip cancer

This means that

\[ p(\theta_i \mid \theta_j \neq i, \xi, y) \propto p(\theta_i) p(\xi_i - \theta_i \mid \{\xi_j - \theta_j \}_{j \neq i}). \]

Since this distribution is free of the data \( y \), the \( \theta_i \) are Bayesianly unidentified (and so are the \( \phi_i \)).

BUT: this does not preclude Bayesian learning about \( \theta_i \); this would instead require

\[ p(\theta_i | y) = p(\theta_i). \]

[Stronger condition: data have no impact on the marginal (not conditional) posterior.]
WinBUGS Example 1: Lip cancer

- **Dilemma:** Though unidentified, the $\theta_i$ and $\phi_i$ are interesting in their own right, as is

\[ \alpha = \frac{sd(\phi)}{sd(\theta) + sd(\phi)}, \]

where $sd(\cdot)$ is the empirical marginal standard deviation.

Are there vague but proper prior values $\tau_h$ and $\tau_c$ that

- lead to acceptable convergence behavior, but
- still allow Bayesian learning?

- Tricky to specify a “fair” prior balance between heterogeneity and clustering (e.g., one for which $\alpha \approx 1/2$) since $\theta_i$ prior is specified marginally while the $\phi_i$ prior is specified conditionally!
WinBUGS Example 1: Lip cancer

★ left panel: $100Y_i/E_i$ (SMR), where $Y_i =$ observed and $E_i =$ expected cases for $I = 56$ districts, 1975–1980

★ right panel: $x_i$, % of the population engaged in agriculture, fishing or forestry (AFF covariate)

★ we also have: a variety of vague, proper, and arguably “fair” priors for $\tau_c$ and $\tau_h$
WinBUGS Example 1: Lip cancer

For actual WinBUGS code, see: http://www.biostat.umn.edu/~brad/data/Lipsbrad.odc

Results:

- AFF covariate appears significantly different from 0 under all 3 priors, although convergence is very slow.
- Excess variability in the data is mostly due to clustering ($E(\alpha|y) > .50$), but the posterior distribution for $\alpha$ does not seem robust to changes in the prior.
- Convergence for the $\xi_i$ (reasonably well-identified) is rapid; convergence for the $\mu_i$ (not shown) is virtually immediate.
WinBUGS Example 1: Lip cancer

Posterior and MCMC convergence summaries:

<table>
<thead>
<tr>
<th>priors for $\tau_c, \tau_h$</th>
<th>posterior for $\alpha$</th>
<th>posterior for $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(1.0, 1.0), G(3.2761, 1.81)$</td>
<td>mean: .57, sd: .058, l1acf: .80</td>
<td>mean: .43, sd: .17, l1acf: .94</td>
</tr>
<tr>
<td>$G(.1, .1), G(.32761, .181)$</td>
<td>mean: .65, sd: .073, l1acf: .89</td>
<td>mean: .41, sd: .14, l1acf: .92</td>
</tr>
<tr>
<td>$G(.1, .1), G(.001, .001)$</td>
<td>mean: .82, sd: .10, l1acf: .98</td>
<td>mean: .38, sd: .13, l1acf: .91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>priors for $\tau_c, \tau_h$</th>
<th>posterior for $\xi_1$</th>
<th>posterior for $\xi_{56}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G(1.0, 1.0), G(3.2761, 1.81)$</td>
<td>mean: .92, sd: .40, l1acf: .33</td>
<td>mean: -.96, sd: .52, l1acf: .12</td>
</tr>
<tr>
<td>$G(.1, .1), G(.32761, .181)$</td>
<td>mean: .89, sd: .36, l1acf: .28</td>
<td>mean: -.79, sd: .41, l1acf: .17</td>
</tr>
<tr>
<td>$G(.1, .1), G(.001, .001)$</td>
<td>mean: .90, sd: .34, l1acf: .31</td>
<td>mean: -.70, sd: .35, l1acf: .21</td>
</tr>
</tbody>
</table>
WinBUGS Example 2: Home prices

Here we illustrate a non-Gaussian model for point-referenced spatial data:

- **Data**: Observations are home values (based on recent real estate sales) at 50 locations in Baton Rouge, Louisiana, USA.
- The response $Y(s)$ is a binary variable, with
  
  $$Y(s) = \begin{cases} 
  1 & \text{if price is “high” (above the median)} \\
  0 & \text{if price is “low” (below the median)} 
  \end{cases}$$

- Observed covariates include the house’s age and total living area
WinBUGS Example 2: Home prices

- We fit a generalized linear model where
  \[ Y(s_i) \sim Bernoulli(p(s_i)), \quad \logit(p(s_i)) = x^T(s_i)\beta + w(s_i) \]

- Assume vague priors for \( \beta \), a Uniform(0, 10) prior for \( \phi \), and an Inverse Gamma(0.1, 0.1) prior for \( \sigma^2 \).

- The WinBUGS code and data for this example are at
  www.biostat.umn.edu/~brad/data/BatonRougebinary.bug:

```winbugs
for (i in 1:N) {
    Y[i] ~ dbern(p[i])
    logit(p[i]) <- w[i]
for (i in 1:3) beta[i] ~ dnorm(0.0,0.001)
w[1:N] ~ spatial.exp(mu[], x[], y[], spat.prec, phi, 1)
phi ~ dunif(0.1,10)
spat.prec ~ dgamma(0.1, 0.1)
sigmasq <- 1/spat.prec
```
WinBUGS Example 2: Home prices

- Use image and contour on \( w_i \) posterior medians in R
- negative residuals (i.e., lower prices) in the north;
  positive residuals (i.e., higher prices) in the south
- smooth flat stretches across the central parts;
  downward slopes toward the north and southeast.
WinBUGS Example 2: Home prices

Parameter estimates (posterior medians and upper and lower .025 points):

<table>
<thead>
<tr>
<th>Parameter</th>
<th>50%</th>
<th>(2.5%, 97.5%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$ (intercept)</td>
<td>$-1.096$</td>
<td>($-4.198, 0.4305$)</td>
</tr>
<tr>
<td>$\beta_2$ (living area)</td>
<td>0.659</td>
<td>($-0.091, 2.254$)</td>
</tr>
<tr>
<td>$\beta_3$ (age)</td>
<td>0.009615</td>
<td>($-0.8653, 0.7235$)</td>
</tr>
<tr>
<td>$\phi$</td>
<td>5.79</td>
<td>(1.236, 9.765)</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>1.38</td>
<td>(0.1821, 6.889)</td>
</tr>
</tbody>
</table>

The covariate effects are generally uninteresting, though living area seems to have a marginally significant effect on price class.