Bayesian Methods

LABORATORY

Lesson 8: March 14 2002

Software: **BUGS**

Multilevel GLM for hierarchical data sets

Multilevel GLM. Small area deaths

Example 8.1 taken from Congdon's book *Poisson model fo* *small area deaths*, p. 272

On the course web page:

- Congdon's BUGS program: *Program 8.1 Multi-level IHD mortality*;
- Extension: **MLpois.b**

A three level hierarchical dataset

Incidence of heart disease mortality in $i = 1, \ldots, 758 =$**I** **small areas** (electoral wards) in the Greater London area over the $3 =$**T years** 1990-92. These small areas are grouped into $j = 1, \ldots, 33 =$**J clusters** (boroughs).

Regressors:

- at the area level: x_{ij}, index of socio-economic deprivation;
- at the cluster level: w_j, 1 for inner London borough, 0 for the outer suburban boroughs.

Theoretical setting

Level-1 model:

$$O_{ij}|\mu_{ij} \overset{ind}{\sim} Pois(\mu_{ij})$$
$$log(\mu_{ij}) = log(E_{ij}) + \beta_{1j} + \beta_{2j}\,(x_{ij} - \bar{x}) + \delta_{ij}$$
$$\delta_{ij} \sim N(0, \sigma^2_\delta)$$

Level-2 model:

$$\boldsymbol{\beta}_j = \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \end{pmatrix} \overset{ind}{\sim} N_2(\boldsymbol{\mu}_{\beta j}, \Omega)$$
$$\mu_{\beta 1j} = \gamma_{11} + \gamma_{12}\, w_j$$
$$\mu_{\beta 2j} = \gamma_{21} + \gamma_{22}\, w_j$$

Priors for Bayesian analysis:

$$\Omega^{-1} \sim W((\rho R)^{-1},\ \rho)$$
$$\gamma_{11} \sim N(0, 1.0E-4)$$
$$\sigma_\delta^{-2} \sim G(a, b)$$

Level-1: the prior specification for $\eta_{ij} = log(\mu_{ij})$

$log(E_{ij})$ is an offset, i.e., an explanatory variable with known coefficient (equal to 1);

$\boldsymbol{\beta} = (\beta_{1j}, \beta_{2j})$ are **random coefficients** for the intercepts and the impacts of deprivation at the cluster level;

δ_{ij} is a random error for Poisson **overdispersion**

$x_{ij} - \bar{x}$ is a area level covariate centred w.r.t. the overall mean.

Level-2: the prior specification for $\boldsymbol{\mu}_{\beta j}$

$\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}$ are the **population coefficients** for the intercept, the impact of borough category, the impact of deprivation, and, respectively, the interaction impact of the level-2 regressor and level-1 regressor.

Alternative likelihood models

1. $O_{ij} \sim Bin(p_{ij}, n_{ij})$, where n_{ij} is the known number of individuals at risk (in area i within cluster j), and p_{ij} is the associated heart disease mortality rate;
2. O_{ij} *or* $\sqrt{O_{ij}}$ *or* $\sqrt{O_{ij}} + \sqrt{O_{ij}+1} \sim Norm(p_{ij}n_{ij}, n_{ij}p_{ij}(1-p_{ij})$, when the counts O_{ij} are sufficiently large.

If the counts consist in small (including even several 0s) values, the choice is preferably

$$O_{ij} \sim Pois(p_{ij}n_{ij}) \\ = Pois(E_{ij}\lambda_{ij})$$

where E_{ij} is the expected count (in area i within cluster j).
In fact, Poisson model is more appropriate than the Binomial one greater the rarity of the studied event is.

The non centered versus the hierarchically centered parameterization

Non-centered parameterization:

$$log(\mu_{ij}) = \gamma_{11} + \gamma_{21}\bar{x}_{ij} + \gamma_{12}w_j + \gamma_{22}\bar{x}_{ij}w_j + \\ \delta_{ij} + \nu_{1j} + \nu_{2j}\bar{x}_{ij}$$

Hierarchically centered:

Level-1 model centered onto:
Level-2 model

Because of non-identifiability issues in GLMMs,
hence high parameter cross-correlation,
it is computationally convenient the centered
parameterization (Gelfand and Sahu, 1999).

Computation of E_{ij}

Suppose that p^* is an overall disease rate. Then,
$E_{ij} = n_{ij}p^*$ and $\lambda_{ij} = p_{ij}/p^*$.
The μs are said

1. externally standardized if p^* is obtained from another data source (such as a standard reference table);
2. internally standardized if p^* is obtained from the given dataset, e.g., $p^* = \sum_{ij} O_{ij} / \sum_{ij} n_{ij}$.

In our example we rely on the latter approach.
Under 1. the joint distribution of the O_{ij} is a product Poisson; under 2. is multinomial.
However, since likelihood inference is unaffected by whether or not we condition on $\sum_{ij} O_{ij}$, the product Poisson likelihood is commonly retained.