Non-Gaussian Case

Consider a distribution in the exponential family. Then, using a canonical parameter representation, we have that

$$p(z_i|z_{-i}) \propto \exp\{\psi(\theta_i z_i - \eta(\theta_i))\}$$

Following the ideas used for the normal CAR models we can set $\theta_i = \sum_{j \neq i}^{n} w_{ij} z_j$, or more generally $\theta_i = x_i' \beta + \sum_{j \neq i}^{n} w_{ij} z_j$. Thus

$$p(z_i|z_{-i}) \propto \exp\left\{ z_i x_i' \gamma + \psi z_i \sum_{j \neq i}^{n} w_{ij} z_j \right\}$$

This model depends on parameters γ and ψ.

An important special case of the previous model is the **autologistic** model. This corresponds to binary variables z_i. Using the logistic link we have that

$$\log \frac{Pr(z_i = 1)}{Pr(z_i = 0)} = x_i'\gamma + \psi \sum_{j \neq i}^{n} w_{ij} z_j$$

Using Brook's lemma we have that

$$p(z_1, \ldots, z_n) \propto \exp \left\{ \gamma' \sum_{i=1}^{n} z_i x_i + \psi \sum_{i,j}^{n} w_{ij} z_j z_i \right\}$$

Unfortunately this joint density is normalized by a constant that depends on γ and ψ. Such constant is practically impossible to compute for large n.

Hierarchical Formulations

A regression model for an m-dimensional vector y of binary variables can be formulated as

$$y_i \sim Ber(g^{-1}(x_i'z)), \quad g(p) = \begin{cases} \log(p/(1-p)) & \text{logit} \\ \Phi(p) & \text{probit} \end{cases}$$

where z is an n-dimensional GMRF, and x_i are covariates.

Hierarchical Formulations

A regression model for an m-dimensional vector y of binary variables can be formulated as

$$y_i \sim Ber(g^{-1}(x_i' z)), \quad g(p) = \begin{cases} \log(p/(1-p)) & \text{logit} \\ \Phi(p) & \text{probit} \end{cases}$$

where z is an n-dimensional GMRF, and x_i are covariates.

These models are equivalent to

$$\varepsilon_i \sim G, \quad \omega_i = x_i' z + \varepsilon_i, \quad y_i = \begin{cases} 1 & \text{if } \omega_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

where G is the normal or logistic distribution. In fact, due to the symmetry of G,
$Pr(y_i = 1) = Pr(\omega_i > 0) = Pr(x_i' z + \varepsilon_i > 0) = G(x_i' z)$.

MCMC

The hierarchical probit CAR depends on three blocks of parameters: z, the latent GMRF; θ, the parameters that control the GMRF; ω, the latent binary variables. We can estimate them using a MCMC. The joint posterior is

$$\pi(z, \omega, \theta | y) \propto \pi(y|\omega)\pi(\omega|z)\pi(z|\theta)\pi(\theta)$$

Suppose $n = m$ and that $x_i' z = z_i$ then

$$\pi(z|\omega, \theta) \propto \exp\left\{-\frac{1}{2}\left(z' Q(\theta) z - \sum_i (z_i - \omega_i)^2\right)\right\}$$

which is proportional to a normal distribution with mean ω and covariance $Q(\theta) + I$, for z. For general covariates we also obtain a normal distribution.

To sample from the posterior $\pi(z, \omega, \theta|y)$ we sample from the block (z, θ) and ω. For ω we have:

$$\pi(\omega|z, y) = \prod_{i=1}^{n} \pi(\omega_i|z_i, y_i)$$

where $\pi(\omega_i|z_i, y_i)$ is truncated normal with mean z_i, variance 1. It is truncated to be positive if $y_1 = 1$ and negative if $y_i = 0$.

To sample from the posterior $\pi(z, \omega, \theta | y)$ we sample from the block (z, θ) and ω. For ω we have:

$$\pi(\omega | z, y) = \prod_{i=1}^{n} \pi(\omega_i | z_i, y_i)$$

where $\pi(\omega_i | z_i, y_i)$ is truncated normal with mean z_i, variance 1. It is truncated to be positive if $y_1 = 1$ and negative if $y_i = 0$.

For (z, θ) we note that

$$\pi(z, \theta | \omega, y) = \pi(z | \omega, \theta) \pi(\theta)$$

so, we can sample θ^* from a proposal distribution. If it is accepted, we then sample z^* from $\pi(z | \omega, \theta^*)$ which is a multivariate normal. This block sampling is more efficient and produces better mixing.

Data on incidence of cervical cancer were obtained for the 216 districts of the former East German Republic in 1979. The cases were classified as premalignant ($y_i = 1$) or malignant ($y_i = 0$). Age is considered as a covariate and is denoted as t.

We use the model

$$\text{probit}(\pi_i) = \alpha + \beta t_i + \gamma_{k(i)}$$

where $k(i)$ denotes the district of the i-th observation. Thus β represents the age effect and $\gamma_{k(i)}$ the district effect.

Focusing in the spatial effects, we have that

$$\gamma_k = u_k + v_k$$

where v_k represents unstructured variation and $v \sim N(0, 1/\kappa_v I)$.
u_k corresponds to the GMRF

$$\pi(u) \propto \kappa_u^{(n-1)/2} \exp\left\{ -\frac{\kappa_u}{2} \sum_{i \sim j} (u_i - u_j)^2 \right\}$$

So we have that

$$\pi(u, v, \omega, \kappa_u, \kappa_v, \beta, \alpha|y) \propto$$

$$\pi(y|\omega)\pi(\omega|\alpha, \beta, u, v)\pi(u|\kappa_u)\pi(v|\kappa_v)\pi(\alpha, \beta)\pi(\kappa_u, \kappa_v)$$

ω is sampled from truncated normal distributions. The joint full conditionals of u and v are multivariate normals. The full conditional of (α, β) is a bivariate normal.

Logistic Link

It is easy to generalize the model for the probit link to any link that depends on a distribution that is a scale mixture of normals. A scale mixture of normals can be written as

$$\pi(x) = \int_{\Lambda} N(0, 1/\lambda)\pi(\lambda)d\lambda$$

for a given density $\pi(\lambda)$. The logistic, the student and the double exponential are common examples of scale mixtures of normals. The mixing distribution of the logistic is the Kolmogorov-Smirnov distribution.

It is easy to generalize the model for the probit link to any link that depends on a distribution that is a scale mixture of normals. A scale mixture of normals can be written as

$$\pi(x) = \int_{\Lambda} N(0, 1/\lambda)\pi(\lambda)d\lambda$$

for a given density $\pi(\lambda)$. The logistic, the student and the double exponential are common examples of scale mixtures of normals. The mixing distribution of the logistic is the Kolmogorov-Smirnov distribution.

Assuming that $\psi_i \sim KS$, then

$$\lambda_i = 1/(2\psi_i)^2, \quad \omega_i \sim N(x_i' z, 1/\lambda_i) \quad y_i = \begin{cases} 1 & \text{if } \omega_i > 0 \\ 0 & \text{otherwise} \end{cases}$$

Suppose $y = (y_1, \ldots, y_n)$ corresponds to the number of deaths from the disease for each of the n counties. We calculate the death rate of the population and then multiply by the population at risk in each county to obtain the expected number of deaths per county e_i.

We assume that y_i has a Poisson distribution with mean $e_i r_i$, where r_i is the **relative risk**. Thus

$$p(y_i | r_i) = \exp\{-e_i r_i\} \frac{(e_i r_i)^{y_i}}{y_i!}$$

here the goal is to estimate r_i.

The MLE of r_i is the **standardized mortality ratio** (SMR) for the i-th area

$$\hat{r}_i = \frac{y_i}{e_i}$$

with estimated standard deviation

$$s_i = \frac{\sqrt{y_i}}{e_i}$$

The MLE of r_i is the **standardized mortality ratio** (SMR) for the i-th area

$$\hat{r}_i = \frac{y_i}{e_i}$$

with estimated standard deviation

$$s_i = \frac{\sqrt{y_i}}{e_i}$$

The SMR does not take into account the population size of the area. So the largest SMRs may correspond to few cases. On the other hand, p -values to compare SMRs to unity are influenced by population size. So the most extreme p-values may simply identify the areas with largest population. These problems are particular important when considering rare diseases over small areas.

Letting $x_i = \log r_i$ we assume that x can be decomposed as

$$x = u + v$$

where v is normal with mean zero and precision matrix $\kappa_v I$. This corresponds to the **unstructured** variability. u is GMRF, so

$$p(u|\kappa_u) \propto \kappa_u^{(n-1)/2} \exp\left\{ -\frac{\kappa_u}{2} \sum_{i \sim j} (u_i - u_j)^2 \right\}$$

Posterior Distribution

The posterior distribution takes the form

$$\pi(u, v, \kappa|y) \propto \kappa_u^{(n-1)/2}\kappa_v^{n/2} \exp\left\{-\frac{\kappa_u}{2}\sum_{i\sim j}(u_i - u_j)^2 - \frac{\kappa_v}{2}\sum_i v_i^2\right\}$$

$$\times \exp\left\{\sum_i y_i(u_i + v_i) - e_i \exp\{u_i + v_i\}\right\} \times \pi(\kappa_v) \times \pi(\kappa_u)$$