Model Summary

Large population case: Poisson data model

Dependent variable:

Count of disease within a small area:

$y_i,\ i = 1, ..., n$

($sometimes\ n_i$ is used instead of y_i)

Each area also has and expected rate/count and a relative risk. These are denoted e_i and θ_i.

Data level model:

$$y_i \sim Pois(\mu_i)$$
$$\mu_i = e_i \theta_i$$
$$\log(\theta_i) = model\ terms$$

Finite population case: binomial model

Note: if you have a count within a finite population in a small area (e. g. birth defects within total births within areas) then the model is naturally binomial. In that case we would have

y_i and n_i

are the case count and population and

p_i the probability of a case in i th area

and

$y_i \sim \text{bin}(p_i, n_i)$

and

$$\text{logit}(p_i) = \log(\frac{p_i}{1-p_i}) = \textit{model terms}$$

Hierarchical Models

Using conditioning we have

$$y \mid \theta$$

$$\theta \mid a, b$$

as a simple model form. Here the data model depends on θ and θ depends on a, b.

Example of a spatial model:

Log-normal model:

$$\log(\theta_i) = \alpha + v_i$$

$$v_i \sim N(0, \tau_v)$$

$$\alpha \sim N(0, \tau_\alpha)$$

Here we have an intercept (α) and an effect in each area that is independent (v_i). They both have normal distributions with zero mean and variance τ_v and τ_α

The BYM model is an extension of this with another effect added:

$$\log(\theta_i) = \alpha + v_i + u_i$$

$$v_i \sim N(0, \tau_v)$$

$$u_i \mid \{u_j\}_{j \neq i} \sim N(\overline{u}_{\delta_i}, \tau_u / n_{\delta_i})$$

$$\alpha \sim N(0, \tau_\alpha)$$

Prior Choice

Prior Choice: some notes

- In Bayesian modeling we usually want to be as ‘non-informative’ as we can be.
- However we can choose priors to ‘fix’ parameters also
- Prior sensitivity is important
- Choice of priors can affect convergence or even run success

Some Recommendations

Regression parameters

- Natural to consider any prior which is zero centered
- And can provide non-informativeness

- Zero mean Gaussian is often used: N(0,tau)
- Double exponential is also used: ddexp(0,tau)
  - Useful in triaging large predictor sets (Machine Learning)

Precisions

- Precisions: usually the SD ~U(0,C)

ie tau<-pow(sd,-2); sd ~dunif(0,C)

- Important when it is a random effect (less so for regression parameter precisions)
- Weakly informative but upper bound must be monitored
- Alternative of tau~Ga(a,b) is used more in INLA and CARBayes, but is less stable computationally in Win/OpenBUGS (see Win/OpenBUGS examples)
- A common weakly informative prior is tau~GA(2,0.5)
- Gamma distribution is conjugate for precisions

Correlation priors

- Correlation is often assigned a uniform prior distribution

$$-1 < \rho < 1$$

$$\rho \sim U(-1,1)$$

or

$$1 < \rho \text{ then}$$

$$\rho \sim U(0,1)$$

$$\text{or } -\log(\rho) \sim Ga(1,1)$$

$$\text{i.e. } -\log(\rho) \sim Exp(1)$$

Probabilities

- Probabilities are on the range (0,1) and
- Often a Beta prior distribution is used (which is conjugate for binomial)

$$p \sim \text{Beta}(1,1)$$

$$\text{or } p \sim \text{Beta}(0.5,0.5)$$

or

$$\text{logit}(p) \sim N(0,\tau)$$