Basic N-mixture Model

Brian M. Brost

25 APR 2015

Model implementation

The file base.N.mixture.sim.R simulates data according to the model statement presented below, and base.N.mixture.mcmc.R contains the MCMC algorithm for model fitting.

Model statement

Let y_{ij} be the j^{th} count of individuals at site i, for $j = 1, \ldots, J$ and $i = 1, \ldots, m$, and N_i be the true number of individuals at site i. Assuming the population is closed to mortality, recruitment, immigration, and emigration over the course of the J surveys conducted at any given site,

$$\begin{aligned}
y_{ij} &\sim \text{Binom}(N_i, p_i) \\
N_i &\sim \text{Pois}(\lambda_i) \\
\lambda_i &\sim \text{Gamma}(r, q) \\
p_i &\sim \text{Beta}(\alpha, \beta).
\end{aligned}$$

Note that this model allows p and λ to vary by site, but these parameters are not modeled as a function of covariates.

Posterior distribution

For a single site, i:

$$[N_i, p_i, \lambda_i | \mathbf{y}_i] \propto \prod_{j=1}^{J} [y_{ij} | N_i, p_i][N_i | \lambda_i][\lambda_i][p_i]$$

Full conditional distributions

Probability of observing an individual (p_i):

$$\begin{aligned}
[p_i | \cdot] &\propto \prod_{j=1}^{J} [y_{ij} | N_i, p_i][p_i] \\
&\propto \prod_{j=1}^{J} p_i^{y_{ij}} (1 - p_i)^{N_i - y_{ij}} p_i^{\alpha - 1} (1 - p_i)^{\beta - 1} \\
&\propto p_i^{\sum_{j=1}^{J} y_{ij}} (1 - p_i)^{\sum_{j=1}^{J} (N_i - y_{ij})} p^{\alpha - 1} (1 - p)^{\beta - 1} \\
&= \text{Beta}\left(\sum_{j=1}^{J} y_{ij} + \alpha, \sum_{j=1}^{J} (N_i - y_{ij}) + \beta \right)
\end{aligned}$$

The true number of individuals (N_i):

$$
\begin{aligned}
[N_i|\cdot] &\propto \prod_{j=1}^{J}[y_{ij}|N_i,p_i][N_i|\lambda_i] \\
&\propto \prod_{j=1}^{J}\left(\begin{array}{c} N_i \\ y_{ij} \end{array}\right) p_i^{y_{ij}}(1-p_i)^{N_i-y_{ij}}\left(\frac{\lambda_i^{N_i}e^{-\lambda_i}}{N_i!}\right) \\
&\propto \prod_{j=1}^{J}\left(\frac{N_i!}{y_{ij}!(N_i-y_{ij})!}\right) p_i^{y_{ij}}(1-p_i)^{N_i-y_{ij}}\left(\frac{\lambda_i^{N_i}e^{-\lambda_i}}{N_i!}\right) \\
&\propto \prod_{j=1}^{J}\frac{(1-p_i)^{N_i-y_{ij}}\lambda_i^{N_i}}{(N_i-y_{ij})!} \\
&\propto \prod_{j=1}^{J}\frac{(1-p_i)^{N_i-y_{ij}}\lambda_i^{N_i}\lambda_i^{-y_{ij}}}{(N_i-y_{ij})!} \\
&\propto \prod_{j=1}^{J}\frac{(\lambda_i(1-p_i))^{\,N_i-y_{ij}}}{(N_i-y_{ij})!}e^{-\lambda_i(1-p_i)}.
\end{aligned}
$$

This full-conditional is a little strange because $[N_i - y_{ij}|\cdot] \propto \text{Pois}(\lambda_i(1-p_i))$, which suggests there is one true abundance per replicate count at each site. This is in contrast to the case in which only one observation exists per site, i.e., $[N_i - y_i|\cdot] \propto \text{Pois}(\lambda_i(1-p_i))$. Given that $[N_i|\cdot]$ lacks a clear analytical solution, sample N_i using Metropolis-Hastings.

Rate of the process model (λ_i):

$$
\begin{aligned}
[\lambda_i|\cdot] &\propto [N_i|\lambda_i][\lambda_i] \\
&\propto \frac{\lambda_i^{N_i}e^{-\lambda_i}}{N_i!}\lambda_i^{r-1}e^{-\lambda_i q} \\
&\propto \lambda_i^{N_i}e^{-\lambda_i}\lambda_i^{r-1}e^{-\lambda_i q} \\
&\propto \lambda_i^{N_i+r-1}e^{-\lambda_i(1+q)} \\
&= \text{Gamma}\left(N_i+r, 1+q\right).
\end{aligned}
$$