

N-MIXTURE MODEL WITH TEMPORAL TREND

Brian M. Brost

25 APR 2015

Model implementation

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

Model statement

Let y_{it} be the i^{th} count of individuals during time period t , for $i = 1, \dots, m$ and $t = 1, \dots, T$, and N_t be the true number of individuals at time t . Assuming the population is closed to mortality, recruitment, immigration, and emigration over the course of the m surveys conducted during any given time period,

$$\begin{aligned} y_{it} &\sim \text{Binom}(N_t, p_{it}) \\ N_t &\sim \text{Pois}(\lambda_t) \\ \lambda_t &= \theta N_{t-1}, t = 2, \dots, T \\ \theta &\sim \text{N}(\boldsymbol{\mu}_\theta, \sigma^2) \\ \lambda_1 &\sim \text{Gamma}(r, q) \\ \text{logit}(p_{it}) &\sim \mathbf{W}_{it}\boldsymbol{\alpha} \\ \boldsymbol{\alpha} &\sim \text{N}(\boldsymbol{\mu}_\alpha, \tau^2 \mathbf{I}). \end{aligned}$$

Note that an identity link for the model on N_t could also be used, i.e., $N_t \sim \text{Pois}(\theta N_{t-1})$. The algorithm in `N.mixture.trend.mcmc.R` will also estimate the posterior predictive distribution for time periods in which no observations are available, i.e., $N_{\tilde{t}}$, $\tilde{t} = 2, \dots, T - 1$.

Posterior distribution

$$[N_1, N_{\{t>2\}}, \boldsymbol{\alpha}, \lambda_1, \theta | \mathbf{Y}, \mathbf{W}] \propto \prod_{t=1}^T \prod_{i=1}^m [y_{it} | N_t, p_{it}] [N_1 | \lambda_1] [N_{\{t>2\}} | \theta] [\theta] [\lambda_1] [\boldsymbol{\alpha}]$$

Full conditional distributions

Coefficients describing the effect of covariates on detection probability ($\boldsymbol{\alpha}$):

$$[\boldsymbol{\alpha} | \cdot] \propto \prod_{t=1}^T \prod_{i=1}^m [y_{it} | N_t, p_{it}] [\boldsymbol{\alpha}].$$

This full-conditional distribution does not have a known analytical form; therefore, sample \mathbf{p} using Metropolis-Hastings.

The true number of individuals during $t = 1$ (N_1):

$$\begin{aligned} [N_1 | \cdot] &\propto \prod_{i=1}^m [y_{i1} | N_1, p_{i1}] [N_1 | \lambda_1] \\ &\propto \prod_{i=1}^m \frac{(\lambda_1 (1 - p_{i1}))^{N_1 - y_{i1}}}{(N_1 - y_{i1})!} e^{-\lambda_1 (1 - p_{i1})}. \end{aligned}$$

This full-conditional is a little strange because $[N_1 - y_{i1}|\cdot] \propto \text{Pois}(\lambda_1(1 - p_{i1}))$, which suggests there is one true abundance per replicate count during $t = 1$. 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 $[N_1|\cdot]$ lacks a clear analytical solution, sample N_1 using Metropolis-Hastings.

Rate of the process model for N_1 (λ_1):

$$\begin{aligned} [\lambda_1|\cdot] &\propto [N_1|\lambda_1][\lambda_1] \\ &\propto \frac{\lambda_1^{N_1} e^{-\lambda_1}}{N_1!} \lambda_1^{r-1} e^{-\lambda_1 q} \\ &\propto \lambda_1^{N_1} e^{-\lambda_1} \lambda_1^{r-1} e^{-\lambda_1 q} \\ &\propto \lambda_1^{N_1+r-1} e^{-\lambda_1(1+q)} \\ &= \text{Gamma}(N_1 + r, 1 + q). \end{aligned}$$

The true number of individuals during $t = 2, \dots, T - 1$ (N_t):

$$[N_t|\cdot] \propto \prod_{i=1}^m [y_{it}|N_t, p_{it}][N_t|\theta, N_{t-1}][N_{t+1}|\theta, N_t]$$

This full-conditional lacks an analytical solution; therefore, sample N_t using Metropolis-Hastings. Note, for time periods \tilde{t} in which observations are missing, the posterior predictive distribution for true abundance is:

$$[N_{\tilde{t}}|\cdot] \propto \prod_{i=1}^m [N_{\tilde{t}}|\theta, N_{\tilde{t}-1}][N_{\tilde{t}+1}|\theta, N_{\tilde{t}}].$$

The true number of individuals during $t = T$ (N_T):

$$[N_T|\cdot] \propto \prod_{i=1}^m [y_{iT}|N_T, p_{iT}][N_T|\theta, N_{T-1}]$$

This full-conditional lacks an analytical solution; therefore, sample N_T using Metropolis-Hastings.

Population growth rate (θ):

$$[\theta|\cdot] \propto [N_t|\theta, N_{t-1}][\theta]$$

This full-conditional lacks an analytical solution; therefore, sample θ using Metropolis-Hastings.