

Multispecies Model

Kristin Broms, Viviana Ruiz-Gutierrez, John Tipton

January 31, 2014

1 Introduction and motivation

2 Model of one season detection

2.1 Modelling assumptions

- There is an ecological motivation for \mathbf{p} and Ψ coming from the same distribution
- There is no heterogeneity in unobserved covariates for \mathbf{p} and Ψ
- All species are "theoretically" available to be sampled
- No seasonality
- Closed populations within the J visits
- The actual number of species N lies between the observed number of species K and the number of augmented species Ω

2.2 Data model

For $i = 1, \dots, n$ sites, $k = 1, \dots, K$ observed species (out of a total of Ω species), and for $j = 1, \dots, J$ revisits at each site

$$y_{ik} = \begin{cases} 0 & \text{if } Z_{ik} = 0 \\ \text{Binom}(J, p_k) & \text{if } Z_{ik} = 1 \end{cases}$$

where y_{ik} is the number of times that species i is seen at site k out of J visits, Z_{ik} is an indicator of presence or absence of species i at site k , and p_k is the detection probability for species k .

2.3 Process

$$Z_{ik} = \begin{cases} 0 & \text{if } W_k = 0 \\ \text{Bern}(\Psi_k) & \text{if } W_k = 1 \end{cases}$$

$$p_k \sim \text{Beta}(\alpha_p, \beta_p)$$

$$\Psi_k \sim \text{Beta}(\alpha_\Psi, \beta_\Psi)$$

$$W_k \sim \text{Bern}(\lambda)$$

where W_k is an indicator of whether an augmented species exists or not in the population and $\sum_{k=1}^{\Omega} W_k = N$, the species richness, Ψ_k is the presence probability for species k , and λ is the probability of an augmented species being real

2.4 Parameter

$$\begin{aligned}
\alpha_p &\sim \text{Gamma}(\alpha_{\alpha_p}, \beta_{\alpha_p}) \\
\beta_p &\sim \text{Gamma}(\alpha_{\beta_p}, \beta_{\beta_p}) \\
\alpha_\Psi &\sim \text{Gamma}(\alpha_{\alpha_\Psi}, \beta_{\alpha_\Psi}) \\
\beta_\Psi &\sim \text{Gamma}(\alpha_{\beta_\Psi}, \beta_{\beta_\Psi}) \\
\lambda &\sim \text{Beta}(\alpha_\lambda, \beta_\lambda)
\end{aligned}$$

3 Ideas

- model $\Psi_k = f(\Psi_{-k}) + \text{randomness and species interaction}$
- $Z_{ik} = \begin{cases} 0 & \text{if } W_k = 0 \\ \text{Bern}(\Psi_k) & \text{if } W_k = 1 \end{cases}$ where W_k represents "ghost species"
- Let Ω represent the total number of possible species, $\Omega \geq N \geq K$
- $W_k \sim \text{Bern}(\lambda) \Leftrightarrow N \sim \text{Binom}(\Omega, \lambda) \Leftrightarrow \sum_{k=1}^{\Omega} W_k = N$
- Goal: estimate Ψ_k and predict N

4 Posterior

$$\begin{aligned}
[\mathbf{Z}, \mathbf{p}, \boldsymbol{\Psi}, \mathbf{W}, \alpha_p, \beta_p, \alpha_\Psi, \beta_\Psi, \lambda | \mathbf{y}, J] &\propto \prod_{i=1}^n \prod_{k=1}^{\Omega} [y_{ik} | J, p_k]^{Z_{ik}} I\{y_{ik} = 0\}^{(1-Z_{ik})} [Z_{ik} | \Psi_k]^{W_k} I\{Z_{ik} = 0\}^{(1-W_k)} \\
&\times [p_k | \alpha_p, \beta_p] [\Psi_k | \alpha_\Psi, \beta_\Psi] [W_k | \lambda] [\alpha_p] [\beta_p] [\alpha_\Psi] [\beta_\Psi] [\lambda]
\end{aligned}$$

5 Full conditionals

5.1 Z_{ik}

For $y_{ik} = 1$, $Z_{ik} = 1$.

For $y_{ik} = 0$ and $W_k = 1$

$$\begin{aligned}
[Z_{ik} | \cdot] &\propto [y_{ik} | J, p_k]^{Z_{ik}} I\{y_{ik} = 0\}^{(1-Z_{ik})} [Z_{ik} | \Psi]^{W_k} I\{Z_{ik} = 0\}^{(1-W_k)} \\
&\propto (1-p_k)^{J Z_{ik}} \Psi_k^{Z_{ik} W_k} (1-\Psi_k)^{(1-Z_{ik}) W_k} \\
&\propto ((1-p_k)^J \Psi_k)^{Z_{ik}} (1-\Psi_k)^{(1-Z_{ik})}
\end{aligned}$$

which is $\text{Bern}(\tilde{\Psi}_k)$ where $\tilde{\Psi}_k = ((1-p_k)^J \Psi_k) / ((1-p_k)^J \Psi_k + (1-\Psi_k))$

For $y_{ik} = 0$ and $W_k = 0$

$$\begin{aligned}
[Z_{ik} | \cdot] &\propto [y_{ik} | J, p_k]^{Z_{ik}} I\{y_{ik} = 0\}^{(1-Z_{ik})} [Z_{ik} | \Psi]^{W_k} I\{Z_{ik} = 0\}^{(1-W_k)} \\
&\propto (1-p_k)^{J Z_{ik}} \Psi_k^{Z_{ik} W_k} (1-\Psi_k)^{(1-Z_{ik}) W_k} \\
&\propto ((1-p_k)^J)^{Z_{ik}}
\end{aligned}$$

which is $\text{Bern}((1 - p_k)^J)$

5.2 p_k

$$\begin{aligned} [p_k|\cdot] &\propto \prod_{i=1}^n [y_{ik}|J, p_k]^{Z_{ik}} [p_k|\alpha_p, \beta_p] \\ &\propto \prod_{i=1}^n (p_k^{y_{ik}} (1 - p_k)^{J - y_{ik}})^{Z_{ik}} p_k^{(\alpha_p - 1)} (1 - p_k)^{(\beta_p - 1)} \\ &\propto p_k^{(\alpha_p - 1 + \sum_{i=1}^n y_{ik} Z_{ik})} (1 - p_k)^{(\beta_p - 1 + \sum_{i=1}^n Z_{ik} (J - y_{ik}))} \end{aligned}$$

which is $\text{Beta}(\alpha_p + \sum_{i=1}^n y_{ik} Z_{ik}, \beta_p + \sum_{i=1}^n Z_{ik} (J - y_{ik}))$

5.3 Ψ_k

When $W_k = 1$

$$\begin{aligned} [\Psi_k|\cdot] &\propto \prod_{i=1}^n [Z_{ik}|\Psi_k]^{W_k} [\Psi_k|\alpha_\Psi, \beta_\Psi] \\ &\propto \prod_{i=1}^n \Psi_k^{Z_{ik}} (1 - \Psi_k)^{(1 - Z_{ik})} \Psi_k^{(\alpha_\Psi - 1)} (1 - \Psi_k)^{(\beta_\Psi - 1)} \\ &\propto \Psi_k^{(\alpha_\Psi - 1 + \sum_{i=1}^n Z_{ik})} (1 - \Psi_k)^{(\beta_\Psi - 1 + \sum_{i=1}^n (1 - Z_{ik}))} \end{aligned}$$

which is $\text{Beta}(\alpha_\Psi + \sum_{i=1}^n Z_{ik}, \beta_\Psi + \sum_{i=1}^n (1 - Z_{ik}))$

When $W_k = 0$

$$\begin{aligned} [\Psi_k|\cdot] &\propto \prod_{i=1}^n [Z_{ik}|\Psi_k]^{W_k} [\Psi_k|\alpha_\Psi, \beta_\Psi] \\ &\propto \prod_{i=1}^n \Psi_k^{Z_{ik}} (1 - \Psi_k)^{(1 - Z_{ik})} \Psi_k^{(\alpha_\Psi - 1)} (1 - \Psi_k)^{(\beta_\Psi - 1)} \\ &\propto \Psi_k^{(\alpha_\Psi - 1)} (1 - \Psi_k)^{(\beta_\Psi - 1)} \end{aligned}$$

which is $\text{Beta}(\alpha_\Psi, \beta_\Psi)$

5.4 W_k

Only sample for $\sum_{i=1}^n y_{ik} = 0$

$$\begin{aligned} [W_k|\cdot] &\propto \prod_{i=1}^n [Z_{ik}|\Psi_k]^{W_k} [W_k|\lambda] \\ &\propto \prod_{i=1}^n \Psi_k^{(Z_{ik} W_k)} (1 - \Psi_k)^{((1 - Z_{ik}) W_k)} \lambda^{W_k} (1 - \lambda)^{(1 - W_k)} \\ &\propto \prod_{i=1}^n \left(\Psi_k^{Z_{ik}} (1 - \Psi_k)^{(1 - Z_{ik})} \lambda \right)^{W_k} (1 - \lambda)^{(1 - W_k)} \end{aligned}$$

which is $\text{Bern}(\tilde{\lambda})$ where $\tilde{\lambda} = \left(\Psi_k^{\sum_{i=1}^n Z_{ik}} (1 - \Psi_k)^{\sum_{i=1}^n (1 - Z_{ik})} \lambda \right) / \left(\Psi_k^{\sum_{i=1}^n Z_{ik}} (1 - \Psi_k)^{\sum_{i=1}^n (1 - Z_{ik})} \lambda + (1 - \lambda) \right)$

5.5 α_p

$$[\alpha_p | \cdot] \propto \prod_{k=1}^{\Omega} [p_k | \alpha_p, \beta_p] [\alpha_p] I\{W_k = 1\}$$

where $[p_k | \alpha_p, \beta_p]$ is $\text{Beta}(\alpha_p, \beta_p)$ and $[\alpha_p]$ is $\text{Gamma}(\alpha_{\alpha_p}, \beta_{\alpha_p})$. This can be sampled using Metropolis-Hastings

5.6 β_p

$$[\beta_p | \cdot] \propto \prod_{k=1}^{\Omega} [p_k | \alpha_p, \beta_p] [\beta_p] I\{W_k = 1\}$$

where $[p_k | \alpha_p, \beta_p]$ is $\text{Beta}(\alpha_p, \beta_p)$ and $[\beta_p]$ is $\text{Gamma}(\alpha_{\beta_p}, \beta_{\beta_p})$. This can be sampled using Metropolis-Hastings

5.7 α_{Ψ}

$$[\alpha_{\Psi} | \cdot] \propto \prod_{k=1}^{\Omega} [\Psi_k | \alpha_{\Psi}, \beta_{\Psi}] [\alpha_{\Psi}] I\{W_k = 1\}$$

where $[\Psi_k | \alpha_{\Psi}, \beta_{\Psi}]$ is $\text{Beta}(\alpha_{\Psi}, \beta_{\Psi})$ and $[\alpha_{\Psi}]$ is $\text{Gamma}(\alpha_{\alpha_{\Psi}}, \beta_{\alpha_{\Psi}})$. This can be sampled using Metropolis-Hastings

5.8 β_{Ψ}

$$[\beta_{\Psi} | \cdot] \propto \prod_{k=1}^{\Omega} [\Psi_k | \alpha_{\Psi}, \beta_{\Psi}] [\beta_{\Psi}] I\{W_k = 1\}$$

where $[\Psi_k | \alpha_{\Psi}, \beta_{\Psi}]$ is $\text{Beta}(\alpha_{\Psi}, \beta_{\Psi})$ and $[\beta_{\Psi}]$ is $\text{Gamma}(\alpha_{\beta_{\Psi}}, \beta_{\beta_{\Psi}})$. This can be sampled using Metropolis-Hastings

5.9 λ

$$\begin{aligned} [\lambda | \cdot] &\propto \prod_{k=1}^{\Omega} [W_k | \lambda] [\lambda] \\ &\propto \prod_{k=1}^{\Omega} \lambda^{W_k} (1 - \lambda)^{(1 - W_k)} \lambda^{(\alpha_{\lambda} - 1)} (1 - \lambda)^{(\beta_{\lambda} - 1)} \end{aligned}$$

which is $\text{Beta}(\alpha_{\lambda} + \sum_{k=K+1}^{\Omega} W_k, \beta_{\lambda} + \sum_{k=K+1}^{\Omega} (1 - W_k))$

6 Next Steps

- Include a probit link function