N Prior Data Augmentation

The most common approach to data augmentation for capture-recapture models (Royle et al., 2007) augments the capture history up to size M and associates indicator variables $z_i \sim \text{Bern}(\psi)$ with individual $i = 1, \ldots, M$, where $N = \sum_{i=1}^{M} z_i$. This approach induces the distribution $N \sim \text{Binomial}(M, \psi)$, and allows N to be estimated by estimating the latent elements of $\boldsymbol{z}$ (all indices except the detected individuals known to be in the population). This approach is often motivated by the observation that as $M \to \infty$, $\psi \to 0$ and the distribution of N converges to a Poisson distribution with parameter $\lambda = M\psi$ (e.g., Royle et al., 2013).

This result is desirable because abundance is a count and is typically modeled as a count random variable; however, when using data augmentation, we would like to set M as low as possible without allowing N to ever reach M to maximize computational efficiency. Thus in real world applications, the Poisson approximation is likely to be poor, and in a multisession model of multiple populations, the variance of N could be underestimated an arbitrary amount depending on where M for each population is set.

A count prior for abundance can be used in capture-recapture models using the complete data likelihood and Reversible Jump MCMC (King and Brooks, 2008; Schofield and Barker, 2008) or data augmentation (Schofield and Barker, 2014) estimation approaches, or using the semi-complete data likelihood (King et al., 2016). For complete data likelihood approaches, Schofield and Barker (2014) distinguish between models where N is a parameter (CDL-N) and those where the indicator variables are parameters. Schofield and Barker (2014) call these CDL-w models because they use w to represent the indicator variables–here we will call these CDL-z since we use z to represent the indicator variables. Schofield and Barker (2014) show how to specify distributions on the individual indicator variables that induce a Poisson distribution on N, and this model can be specified with BUGS code and fit in programs like JAGS (Plummer, 2025) and NIMBLE (de Valpine et al., 2017). Schofield and Barker (2014) also describe how to fit CDL-N using the algorithm of Carlin and Chib (1995) and provide BUGS code allowing it to be fit in JAGS or NIMBLE. Here, we provide a modification of this approach that is generally more flexible and computationally efficient.

First, here is the joint posterior for a typical SCR model replacing the individual Bernoulli priors on z_i that induce a Binomial prior on N with a Poisson prior directly on N:

$$[N, \boldsymbol{z}, \boldsymbol{s}, \lambda, p_0, \sigma | \boldsymbol{Y}, \boldsymbol{X}] \propto \left\{ \prod_{i}^{M} \left\{ \prod_{j}^{J} \left\{ \prod_{k}^{K} [y_{i,j,k} | z_i, \boldsymbol{s}_i, \sigma, p_0] \right\} \right\} [\boldsymbol{s}_i] \right\} [\boldsymbol{z}|N][N][\lambda][p_0][\sigma] \quad \textbf{(1)}$$

where $N \sim$ Poisson(λ), λ is the expected abundance, $\boldsymbol{s}_i$ is the activity center of individual i, p_0 is the baseline detection probability, σ is the spatial scale parameter, and $y_{i,j,k}$ is the detection outcome of individual i at trap j on occasion k. Schofield and Barker (2014) assume that the capture histories are ordered with the individuals truly in the population at the top and the "false" individuals at the bottom. Then, when updating N, individuals are added or subtracted in order from the current value of N. This approach requires that

$$P(\boldsymbol{z}|N) = \frac{1}{\binom{M}{N}(N-n)!(M-N)!} \text{ if } \sum_{i=1}^{M} z_i = N, \quad \text{else } 0, \quad \textbf{(2)}$$

where $\binom{M}{N}$ is the number of ways to choose N real individuals from M possible individuals, $(N-n)!$ is the number of ways to order the uncaptured, true individuals and $(M-N)!$ is the number of ways to order the false individuals.

Using the data augmentation approach of Royle et al. (2007) individuals are not assumed to be ordered, allowing any z_i to be turned on or off regardless of preceding or subsequent values of $\boldsymbol{z}$. This feature is especially convenient in latent identity SCR models where the assignment of samples to individuals changes on each MCMC iteration and the individuals assigned detections do not remain at the top of the capture history (e.g., Chandler and Royle, 2013). Therefore, we will assume that

$$P(\boldsymbol{z}|N) = \frac{1}{\binom{M}{N}} \text{ if } \sum_{i=1}^{M} z_i = N, \quad \text{else } 0. \quad \textbf{(3)}$$

Using this modification from Schofield and Barker (2014) no longer allows updates of N and $\boldsymbol{z}$ to be specified using the BUGS code alone–we describe a custom update below implemented in NIMBLE.

Metropolis-Hastings Algorithm

(0) Initialize all parameters, ensuring $\sum_{i=1}^{M} z_i^{(0)} = N^{(0)}$. We will not include any formal constraints that maintain this relationship, relying on it to be met in the initialized data and respected in the proposal algorithm.

(1) Propose to add or subtract 1 from N^{t-1} with probability η. We will set η to 0.5 so this component of the proposal algorithm is symmetric. If we propose to add, set $N^* = N^{t-1} + 1$, otherwise, set $N^* = N^{t-1} - 1$. If we propose to add, select an index, i^* of $\boldsymbol{z}$ from the set of indices for which $z_i = 0$, which is of size $M - N^{t-1}$. Make this selection with a uniform probability. If we propose to subtract, select an index, i^* of $\boldsymbol{z}$ from the set of indices for which $z_i = 1$, which is of size N^{t-1}.

(2) Compute the acceptance probability $\alpha(x^{(t)}, x^*)$ where $x^{(t)}$ is the current state of N and $\boldsymbol{z}$ and x^* is their proposed state. The acceptance probability is:

$$\alpha(x^{(t)}, x^*) = min\left(1, \frac{\pi(x^*)q(x^{(t)}|x^*)}{\pi(x^{(t)})q(x^*|x^{(t)})}\right) \quad \textbf{(4)}$$

The target distributions is $\pi(x) = [N][\boldsymbol{z}|N][\boldsymbol{Y}_{i^*}|\boldsymbol{z}_{i^*}, \boldsymbol{s}_{i^*}, p_0, \sigma]$. If we are adding, $q(x^{(t)}|x^*) = \frac{\eta}{M-N^{(t)}}$, and if we are subtracting, $q(x^{(t)}|x^*) = \frac{\eta}{N^{(t)}}$. Note, that the proposal ratio and the prior ratio of $[\boldsymbol{z}|N]$ cancel. When adding, $N^* = N^{(t)} + 1$ and the proposal ratio is:

$$\frac{q(\text{add})}{q(\text{subtract})} = \frac{1/(N^{(t)}+1)}{1/(M-N^{(t)})} = \frac{M-N^{(t)}}{N^{(t)}+1}. \quad \textbf{(5)}$$

Then, the prior ratio of $[\boldsymbol{z}|N]$ is:

$$\frac{[\boldsymbol{z}^{(t)}|N^{(t)}]}{[\boldsymbol{z}^{(*)}|N^{(*)}]} = \frac{\frac{1}{\binom{M}{N^{(t)}}}}{\frac{1}{\binom{M}{N^{(t)}+1}}} = \frac{\binom{M}{N^{(t)}+1}}{\binom{M}{N^{(t)}}} = \frac{N^{(t)}+1}{M-N^{(t)}}. \quad \textbf{(6)}$$

which cancel in the Metropolis-Hastings Ratio. The same is true for the subtract step.

(3) Repeat steps 1 and 2 an arbitrary number of times per iteration. We found that updating 25% of M on every occasion to be a generally good choice for balancing the mixing of the MCMC chains and computational efficiency.

Metropolis-Hastings algorithm - alternate proposal

The algorithm above is convenient because the z priors do not need to be computed, but it also allows for the selection of individuals who cannot be turned off–the individuals who were detected. In this case, we can autoreject once they are selected. Alternatively, we can exclude them from selection by restricting the selectable individuals in the subtract step to the $N - n\ z_i = 1$ individuals assigned samples. For regular SCR, these will be indices larger than n^{cap}, the number captured, and for latent ID SCR, these are individuals who are not currently assigned detections. When using this approach, we must use the z priors and proposal probabilities. When adding, the proposal probability is now:

$$\frac{q(\text{add})}{q(\text{subtract})} = \frac{1/(N^{(t)} + 1 - n^{cap})}{1/(M - N^{(t)})} = \frac{M - N^{(t)}}{N^{(t)} + 1 - n^{cap}}. \quad \textbf{(7)}$$

and, canceling terms, the product of the prior and proposal ratio is

$$\frac{N^{(t)} + 1}{N^{(t)} + 1 - n^{cap}}. \quad \textbf{(8)}$$

Similarly, in the subtract proposal, the product of the prior and proposal ratio is:

$$\frac{N^{(t)} - n^{cap}}{N^{(t)}}. \quad \textbf{(9)}$$

A third option is that multiple indices of z may be updated at a time, but we do not describe that here.

References

Carlin, B. P. and Chib, S. Bayesian model choice via markov chain monte carlo methods. *Journal Of The Royal Statistical Society Series B: Statistical Methodology*, 57(3):473–484, 1995.

Chandler, R. B. and Royle, J. A. Spatially explicit models for inference about density in unmarked or partially marked populatioins. *The Annals of Applied Statistics*, pages 936–954, 2013.

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Temple Lang, D., and Bodik, R. Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26:403–413, 2017. doi: 10.1080/10618600.2016.1172487.

King, R. and Brooks, S. On the bayesian estimation of a closed population size in the presence of heterogeneity and model uncertainty. *Biometrics*, 64(3):816–824, 2008.

King, R., McClintock, B. T., Kidney, D., and Borchers, D. Capture–recapture abundance estimation using a semi-complete data likelihood approach. 2016.

Plummer, M. *rjags: Bayesian Graphical Models using MCMC*, 2025. URL https://CRAN.R-project.org/package=rjags. R package version 4-17.

Royle, J. A., Dorazio, R. M., and Link, W. A. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics*, 16(1):67–85, 2007.

Royle, J. A., Chandler, R. B., Sollmann, R., and Gardner, B. *Spatial capture-recapture*. Academic press, 2013.

Schofield, M. R. and Barker, R. J. A unified capture-recapture framework. *Journal of agricultural, biological, and environmental statistics*, 13(4):458–477, 2008.

Schofield, M. R. and Barker, R. J. Hierarchical modeling of abundance in closed population capture–recapture models under heterogeneity. *Environmental and Ecological Statistics*, 21(3):435–451, 2014.