

A model of absolute fossil occurrence rate

Peter D Smits
Committee on Evolutionary Biology, University of Chicago

February 17, 2015

1 Methods

1.1 Fossil occurrence information

Foote and Miller data.

1.2 Hierarchical counting model

1.2.1 Likelihood

First, define y_i as some count of fossil occurrences of genus j in a geologic stage for $i = 1, \dots, n$ and $j = 1, \dots, J$.

The Poisson distribution is commonly used to model count data such as the number of observed fossils. The Poisson distribution has one parameter λ which is a rate or inverse-scale parameter. λ can be interpreted as the expected count observed $\mathbf{E}[y]$. λ can be reparameterized as a regression using the log link function $\mathbf{E}[y] = \exp(\alpha)$ where α can be any real number. This is written formally as

$$\begin{aligned} y_i &\sim \text{Poisson}(\lambda_i) \\ \lambda_i &= \exp(\alpha_i). \end{aligned} \tag{1}$$

Currently, this model (Eq. 1) does not take into account the generic membership j of the fossil count and assumes that all genera have the same sighting rate. To account for variation in occurrence rate between genera while also modeling mean generic occurrence rate I take a Bayesian hierarchical modeling approach (?). First, I redefine α_i as $\alpha_{j[i]}$ to indicate that observation i is a member of genus j . I then assume that genera can be considered exchangeable or that the

actual value of j has no meaning. Given this assumption, values of $\alpha_{j[i]}$ are given the same, shared prior written as

$$\alpha_{j[i]} \sim \mathcal{N}(0, \sigma_j). \quad (2)$$

The scale hyperparameter of the shared prior (Eq. 2) is itself estimated from the data itself. This approach allows genera with small sample size to pull towards the mean of the prior while still genera with large sample sizes and strong effects to be modeled.

This hierarchical approach can be further extended to allow for genus ordinal membership. Define k as the order that genus j belongs to, where $k = 1, \dots, K$. Then, instead of assuming the mean of the prior on α is 0 (Eq. 2), and instead have the mean of the prior be $\mu_{k[j]}$ which is the estimate of the rate of fossil occurrence for order k . Then, assuming that orders are exchangeable, values of $\mu_{k[j]}$ are given the same, shared prior. These changes are then written as

$$\begin{aligned} \alpha_{j[i]} &\sim \mathcal{N}(\mu_{k[j]}, \sigma_j) \\ \mu_{k[j]} &\sim \mathcal{N}(0, \sigma_k). \end{aligned} \quad (3)$$

The current model (Eq. 1) does not take into account the number of chances to count an observation. For example, if counting the number of traffic accidents at a street corner it matters if 20 vehicles have passed through the intersection versus 100. To account for this we can define an exposure term u_i for each observation. In this study, u_i is defined as the number of localities species i occurred in during the given stage. The inclusion of u_i is formulated as

$$\begin{aligned} y_i &= \text{Poisson}(u_i \lambda_i) \\ \lambda_i &= \exp(\log(u_i) + \alpha_{j[i]}). \end{aligned} \quad (4)$$

The inclusion of $\log(u_i)$ in the parameterization of λ_i (Eq. 4) is due to the following relationships

$$\begin{aligned} \frac{\mathbf{E}[y_i]}{u_i} &= \lambda_i \\ \mathbf{E}[y_i] &= u_i \lambda_i \\ \log(\mathbf{E}[y_i]) &= \log(u_i) + \log(\lambda_i) \end{aligned}$$

We can now interpret λ as the expected number of co-occurring species per locality for a given observation. While u_i is called the exposure, $\log(u_i)$ is called the offset (?).

One of the major assumptions of the Poisson distribution is that, because there is only one parameter, the variance of the distribution is equal to the mean ($\frac{\text{Var}[y]}{\text{E}[y]}$). When variance is greater than the mean, this is called overdispersion. We can relax this assumption by assuming that, instead of a Poisson distribution,

observations are drawn from a negative binomial distribution (?). Here, I use the following parameterization of the negative binomial

$$\text{Negative binomial}(y|\eta, \phi) = \binom{y + \phi - 1}{y} \left(\frac{\eta}{\eta + \phi} \right)^y \left(\frac{\phi}{\eta + \phi} \right)^\phi \quad (5)$$

where η is the mean and ϕ is the overdispersion. Substituting the negative binomial for the Poisson, the aspects model current defined are written

$$\begin{aligned} y_i &= \text{Negative binomial}(u_i \eta_i, \phi_y) \\ \eta_i &= \exp(\alpha_{j[i]}) \\ \alpha_{j[i]} &\sim \text{Normal}(\mu_{k[j]}, \sigma_j) \\ \mu_{k[j]} &\sim \text{Normal}(0, \sigma_k). \end{aligned} \quad (6)$$

Finally, given the Bayesian framework taken here, there is the assignment of priors to various non-hierarchically modeled parameters. Scale parameters were given weakly informative half-Cauchy (C^+) priors because they have good regulatory priors for constraining hierarchical effects (??). The priors used here are

$$\begin{aligned} \phi_y &\sim C^+(2.5) \\ \sigma_j &\sim C^+(2.5) \\ \sigma_k &\sim C^+(2.5). \end{aligned}$$

The Cauchy distribution is equivalent to the t -distribution with 1 degree of freedom, and the half-Cauchy distribution is the Cauchy folded about 0.