

ESTIMATING VOCAL REPERTOIRES FROM LIMITED SAMPLES

RICHARD MCELREATH

1. PURPOSE

Suppose a population of N individuals (monkeys, people) who produce among them M different vocalizations. Each individual has potentially a different repertoire of these M vocalizations. We want to characterize repertoire at the individual level, so that we might study its relationship with development for example.

One problem is the undercounting of repertoire which is expected in any finite sample. If some vocalizations are rare, and recordings are short or small in number, then on average we will undercount true repertoire size.

2. PRIOR WORK

This problem has been studied at least since Wildenthal (1965), which fit an exponential curve to sequence (recording) count. Garamszegi et al. (2004) repurposed capture-recapture models for this purpose, which is a more principled idea. Kershenbaum et al. (2015) build a probabilistic model that allows for different vocalization probabilities and model the total repertoire size.

The approach developed below is most similar to Kershenbaum et al. (2015), but allows for variable durations of recording (sampling effort) and for each vocalization to have its own rate of production and probability of being possessed by a single individual. The approach here also computes not only estimated repertoire size, but posterior probabilities for each component of the repertoire (that was not observed). Since this approach uses the Stan language, it is also easily extended to allow for stratification by age, season or other covariates.

3. MODELING STRATEGY

We seek a vector ϕ of posterior probability distributions of length M for each individual i . In the case that i was observed to produce a m , then ϕ_m is 1. But when m was not observed, we can compute the posterior probability that this was due to the finite sample, an undercount. In the simplest model, assume that each vocalization m is produced at a rate λ_m , when an individual possesses it. This produces a Poisson

DEPARTMENT OF HUMAN BEHAVIOR, ECOLOGY AND CULTURE, MAX PLANCK INSTITUTE FOR EVOLUTIONARY ANTHROPOLOGY, DEUTSCHER PLATZ 6, 04103 LEIPZIG, GERMANY

E-mail address: richard_mcelreath@eva.mpg.de.

Date: May 31, 2023.

distribution of the counts $Y_{i,j,m}$ of each possessed vocalization m in a focal recording j of duration d_j .

$$Y_{i,j,m} \sim \text{Poisson}(\lambda_m d_j x_{i,m})$$

30 where $x_{i,m}$ is an indicator variable for whether individual i possesses vocalization m . When the observed $y = Y_{i,j,m} > 0$, then the probability is:

$$\Pr(y|y > 0) = p_m \frac{(\lambda_m d_j)^y \exp(-\lambda_m d_j)}{y!}$$

where p_m is the probability an individual possesses m . When $y = 0$ is observed, we require a mixture:

$$\Pr(y|y = 0) = p_m \exp(-\lambda_m d_j) + (1 - p_m)(1)$$

35 Let $X_{i,m}$ be the total number of times i was observed to use m and D_i be the total sampling effort (duration of recordings). If $X_{i,m} > 0$, then i possesses the vocalization. But what about when $X_{i,m} = 0$? In that case, the posterior probability that individual i possesses vocalization m follows from Bayes' rule:

$$\begin{aligned} \Pr(x_{i,m} = 1|X_{i,m} = 0, D_i) &= \frac{\Pr(0|x_{i,m} = 1) \Pr(x_{i,m} = 1)}{\Pr(0)} \\ &= \frac{\Pr(0|\lambda_m D_i) p_m}{\Pr(0|\lambda_m D_i) p_m + (1 - p_m)} \end{aligned}$$

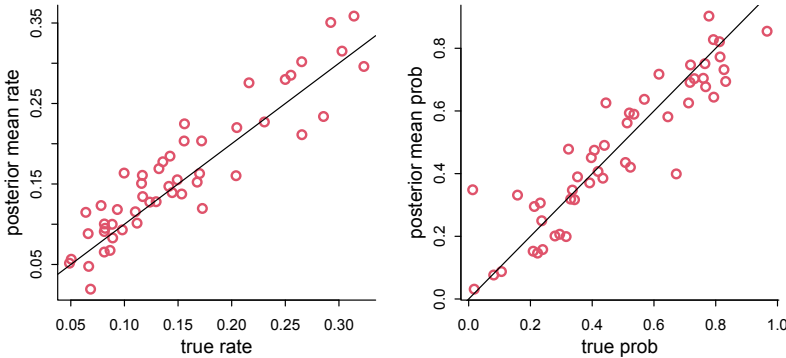
And this is sufficient to solve the problem.

40 In the Poisson case, we need only the total duration of the recording. This model can be modified to allow for Markov transitions between vocalizations, but then we require time stamps for each vocalization.

4. IMPLEMENTATION

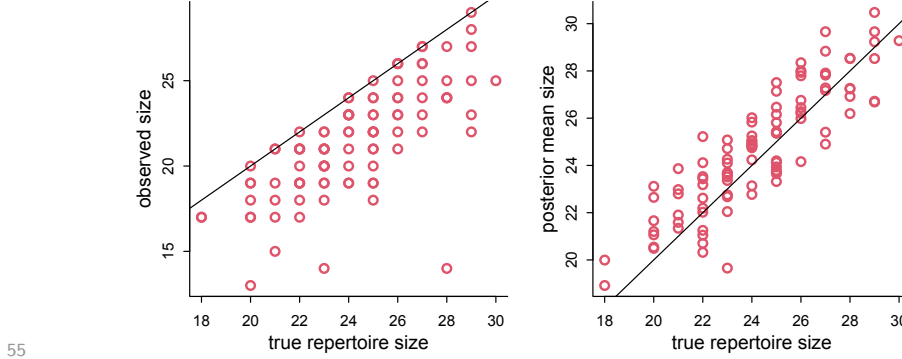
45 We developed a generative model of the sampling design, in which each individual has a variable number of recordings of variable length. Then we used Stan to sample from the posterior distribution of the inferred individual repertoires.

First we can ask how well the model recovers the rate λ and probability p variables.



50 The left compares the true λ values (horizontal) to the posterior means. The diagonal line shows identity. The right compares the true p parameters (horizontal) to the posterior means. This is only one example, but the model unsurprisingly has no trouble estimating Poisson rates and the base rates of each vocalization, except in cases in which a vocalization is never performed in a particular sample.

Now we can ask how the individual repertoires are captured.



The left compares the true repertoire size of each individual (horizontal) to the raw observed (vertical). Note the undercounting. The right compares the same true repertoires (horizontal) to the posterior mean sizes. These are still uncertain, but less biased.

60

5. ADDING AGE

As an example of extending the model with covariates, now we modify the probabilities p so they are an increasing function of individual age. This implies that older individuals possess on average more unique vocalizations. A simple model says that the probability an individual of age a possesses vocalization m is:

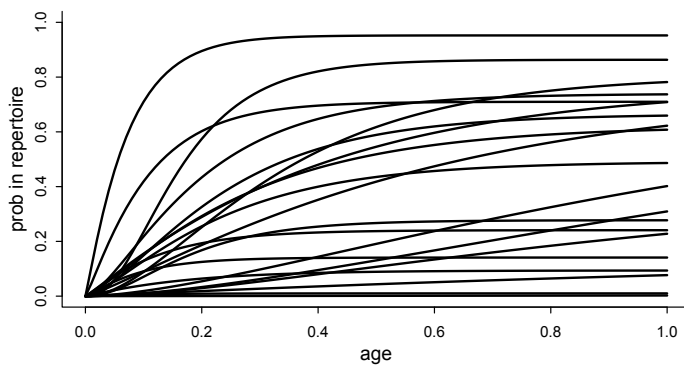
$$p_m(a) = \alpha_m(1 - \exp(-\beta_m a))^{\gamma_m}$$

65 where α_m is the maximum probability for m , β_m is the rate of increase with age, and γ_m is the “elasticity” with age (which allows for slow initial increases when $\gamma_m > 1$).

To see how flexible this function is, suppose we sample functions from the prior:

$$\alpha_m \sim \text{Beta}(2, 2) \quad \beta_m \sim \text{Exponential}(0.25) \quad \gamma_m \sim 1 + \text{Exponential}(2)$$

Here are 20 functions sampled from this prior:



- 70 However we do not argue this function is ideal or even correct. It is just useful for testing. Generalized splines or Gaussian processes could be used instead, if one does not want to impose monotonicity.

REFERENCES

- 75 Garamszegi, L. Z., Balsby, T. J. S., Bell, B. D., Borowiec, M., Byers, B. E.,
Draganoiu, T., Eens, M., Forstmeier, W., Galeotti, P., Gil, D., Gorissen, L.,
Hansen, P., Lampe, H. M., Leitner, S., Lontkowski, J., Nagle, L., Nemeth, E.,
Pinxten, R., Rossi, J.-M., Saino, N., Tanvez, A., Titus, R., Török, J., Duyse,
80 E. V., and Møller, A. P. (2004). Estimating the complexity of bird song by using
capture-recapture approaches from community ecology. *Behavioral Ecology and
Sociobiology*, 57(4):305–317.
- Kershenbaum, A., Freeberg, T. M., and Gammon, D. E. (2015). Estimating vocal
repertoire size is like collecting coupons: A theoretical framework with hetero-
geneity in signal abundance. *Journal of Theoretical Biology*, 373:1–11.
- 85 Wildenthal, J. L. (1965). Structure in primary song of the mockingbird (*mimus
polyglottos*). *The Auk*, 82(2):161–189.