Supplement: A four-step Bayesian workflow for improving ecological science

EM Wolkovich, TJ Davies, WD Pearse & M Betancourt

April 7, 2024

A brief review of statistical inference using Bayesian approaches

Robust analyses ely on our inferences being consistent with the underlying truth more often than not. Quantifying this consistency is calibration—analyzing how often a parameter estimate is close to the true value—a critical part of using models for inference. A major problem with traditional (frequentist) approaches in ecology today is their inferences are unpredictable when their foundational assumptions are not met, but ecologists are not usually trained in how to recognize or deal with this.

To better understand our workflow, we provide a very brief overview of some of the fundamentals of Bayesian methods that is inherently incomplete and, by design, not very technical. This section can be skipped for those who feel already well-versed, and can be augmented for those who are new to Bayesian approaches (for example, MacElreath, 2016; Gelman *et al.*, 2014; Gelman & Hill, 2020).

Probability is often defined as "the long-term frequency with which something happens." We would expect, for example that if we tossed a coin 100 times we would see roughly 50 heads. In this case we would say that the probability of tossing a coin and getting a 'head' is $\frac{50}{100}$, equivalent to 50% or $\frac{1}{2}$. At the same time we wouldn't be very surprised if we observed 49 or even 55 heads, although we would be surprised if we saw 99.

This definition of probability—which is the *frequentist* definition—is useful in many situations, but it has a few disadvantages. First, frequentist definitions aren't very helpful when dealing with unexpected situations. Frequentist probabilities are grounded in repeatable observations, and so understanding these repeatable frequencies is of limited use when trying to make predictions for changing or entirely novel systems. Second, most common frequentist approaches rely on specific assumptions. Using frequentist statistics requires trying to match a model that we have (often just in our heads) of some ecological system to a frequentist method that mostly closely matches the assumptions of our biological model. Given the complexity of ecological data and our uncertainty about the underlying model, frequentist approaches can be especially challenging in ecology.

Bayesian approaches provide a way to build models that can propagate our (un)certainties about what we do—or especially, don't—know about how ecological systems work. In the Bayesian view,

probability is used to quantify uncertainty: the higher the probability of a certain interval of values the less uncertainty we have that the true behavior falls within that interval. Assuming we also have some estimate of our initial uncertainty—usually from knowledge of ecology—to inform a prior distribution (termed below *prior*), then we can apply Bayes' theorem:

$$probability = \frac{likelihood \cdot prior}{normaliser} \tag{1}$$

to update that prior into a posterior distribution that accounts for the information added by a likelihood (*likelihood* is the same as a frequentist likelihood). Here, `normaliser` is a mathematical constant that makes sure our probability cannot go above 100% or below 0% (statisticians are lazy, and will not 'give 110%'). This mathematical constant [technically it is the probability of our data; $p(data)$] is a nuisance term that is extremely challenging to estimate (sometimes it is impossible!) and held back the practical use of Bayesian statistics for almost a century because that normaliser could rarely be analytically worked out. But one of the major advantages of Bayesian methods is that the solution to this problem—numerical simulations based on Markov Chain Monte Carlo (MCMC)—provide a huge additional advantage. Now that no analytical solution needs be found, any model that can be written out mathematically can be fit to data, giving the scientist more freedom of model structure.

Which workflow?

Formally, all a 'workflow' does is organize various steps together in a systematic fashion, but there are many different workflows depending on what the aim is, which will determine which steps a workflow should include. For example a workflow aimed at calibration could look like an expanded version of our Step 1, where all the steps focus on investigating the assumptions encoded in a given model using simulated data. Or a workflow aimed at inference could expand Step 3, to focus on constructing a posterior, then investigating its model adequacy via several criteria. An inferential workflow can also be extended into a model development workflow. If the model adequacy criteria inform not only that something is inadequate about the current model assumptions but what is inadequate (ideally this happens some in Step 4) then one can use those hints to iterative improve the modeling assumptions. We present in the main text a very simplified model development workflow that combines calibration, inference and some model development, but it is not necessarily appropriate for everyone, depending on their aims.

An example workflow

Our workflow in the main text is explained mostly program-agnostically. Though at times we assume a user of Stan, a relatively new probabalistic programming language, that interfaces with R, Python, Julia (and more) to write bespoke Bayesian models (Carpenter *et al.*, 2017). We focus on Stan as its MCMC algorithm (a variant of Hamiltonian Monte Carlo, HMC) is fast and produces specific output to warn of model fit issues (i.e., divergent transitions) in a way other MCMC algorithms do not (e.g. Metropolis-Hastings or Gibbs), but the basic workflow should apply to diverse

implementations of Bayesian modeling, and can be extended to other approaches (frequentist, resampling, etc.).

For our example workflow we use Stan and R. For now, you can download the workflow here: `https://github.com/lizzieinvancouver/bayesianflows/blob/main/examples/synchrony/example.html`.

References

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P. & Allen, R. (2017) Stan: A probabilistic programming language. *Journal of Statistical Software* **76**, 10.18637/jss.v076.i01.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2014) *Bayesian Data Analysis*. CRC Press, New York, 3rd edn.

Gelman, A. & Hill, J. (2020) *Regression and Other Stories*. Cambridge University Press.

MacElreath, R. (2016) *Statistical Rethinking*, vol. 469 pp. CRC Press, New York.