Bayesian Synthetic Likelihood

Thanasi Bakis, Brian Schetzsle

March 16, 2022

Original Paper(s)

- Wood, S. N. (2010), “Statistical Inference for Noisy Nonlinear Ecological Dynamic Systems,” *Nature*.
  - Introduced *Synthetic Likelihood* technique
- L. F. Price, C. C. Drovandi, A. Lee & D. J. Nott (2018), “Bayesian Synthetic Likelihood”, *Journal of Computational and Graphical Statistics*.
  - Extended to *Synthetic Likelihood* technique using standard Bayesian methodology

small history of SL, why it came up in the wood paper & frequentist inference

how it can be easily brought to the bayesian setting with a prior

Synthetic Likelihood and Bayesian Synthetic Likelihood, An Overview

Synthetic likelihood uses a multivariate normal distribution to approximate the density of a vector of summary statistics. This technique is useful in situations where inference on model parameters is desired but calculating model likelihood is difficult or impossible, in which case standard maximum likelihood estimation is not available. If the model is easy to sample from, then different combinations of parameters can be used to generate data which is then condensed into a vector of summary statistics and compared to that of the observed data. Parameter combinations that yield statistics which resembles the observed statistics are deemed more likely. The summary statistics are chosen to capture relevant features of the model and elide problematic features such as noise.

overview of the method itself

IMAGE

Bayesian Synthetic Likelihood, a Toy Example

Model:

$$Y_i \sim \textit{Poisson}(\lambda = 30) \qquad i = 1, ..., 100$$

$$\lambda \sim \textit{Gamma}(\alpha = 0.001, \beta = 0.001)$$

We want to find $P(\lambda|Y) \propto P(Y|\lambda)P(\lambda)$ without evaluating $P(Y|\lambda)$.

Bayesian Synthetic Likelihood, a Toy Example

Note:

$$P(\lambda|Y) \propto \left[\prod_{i=1}^{100} \frac{\lambda^{y_i}}{y_i!} e^{-\lambda}\right] \left[\frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\right]$$

$$\propto \lambda^{\alpha+\sum_{i=1}^{100} y_i - 1} e^{-\lambda(\beta+n)}$$

$$\sim Gamma(\alpha = 0.001 + \sum_{i=1}^{100} y_i, \beta = 100.001)$$

So in this toy example the posterior distribution is known analytically.

Bayesian Synthetic Likelihood, a Toy Example

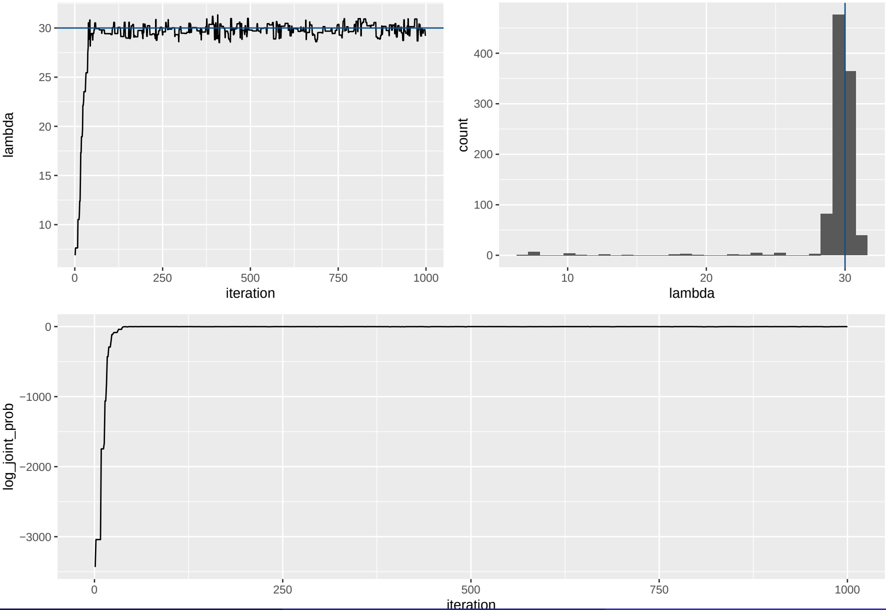
Choosing a statistic:

The paper uses the mean as the statistic to summarize both the observed and generated data:

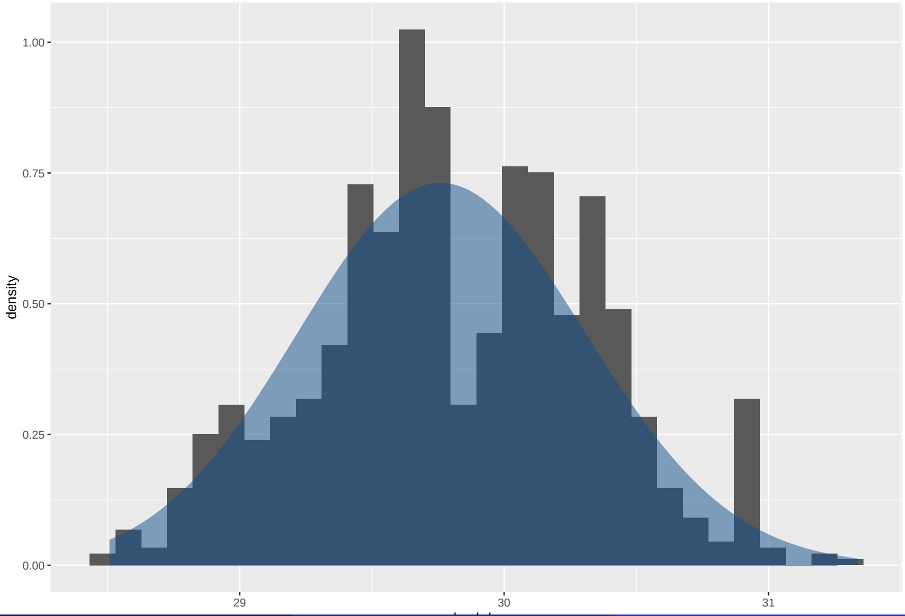
$$s_Y = \frac{1}{100} \sum_{i=1}^{100} Y_i$$

This is the sufficient statistic for the Poisson distribution; all the information contained in the data is also contained in this statistic. Also, by the central limit theorem, the distribution of the mean of a Poisson sample can be adequately approximated by a normal distribution, so synthetic likelihood should perform well in this setting.

Bayesian Synthetic Likelihood, a Toy Example



Bayesian Synthetic Likelihood, a Toy Example



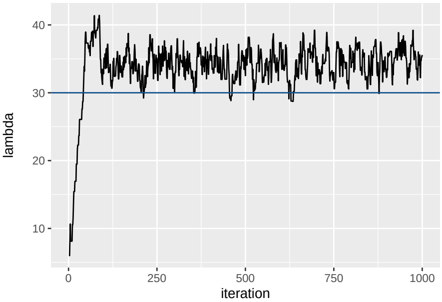
Bayesian Synthetic Likelihood, Further Exploration

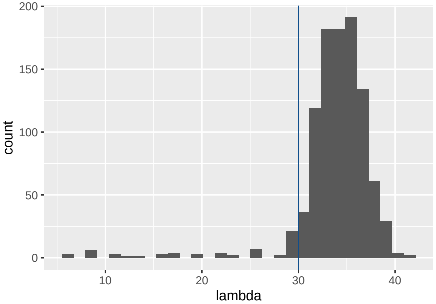
We were interested in trying other statistics to see how well synthetic likelihood performed. This was our own exploration and was not addressed in the paper. We started with the maximal statistic

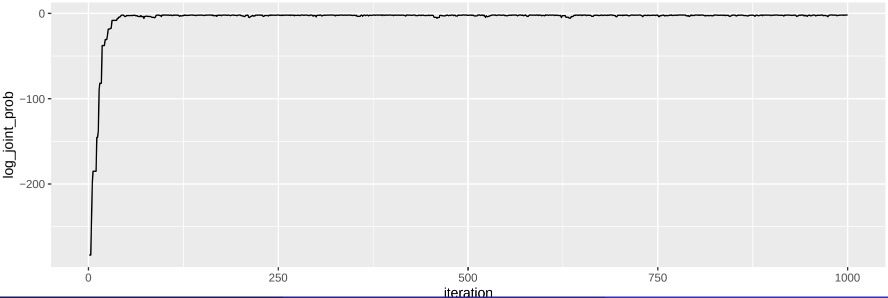
$$s_Y = \max(Y)$$

This is not a sufficient statistic for Poisson data and also not approximately normally distributed across many samples. We would expect the synthetic likelihood method to have a harder time identifying the true analytic posterior.

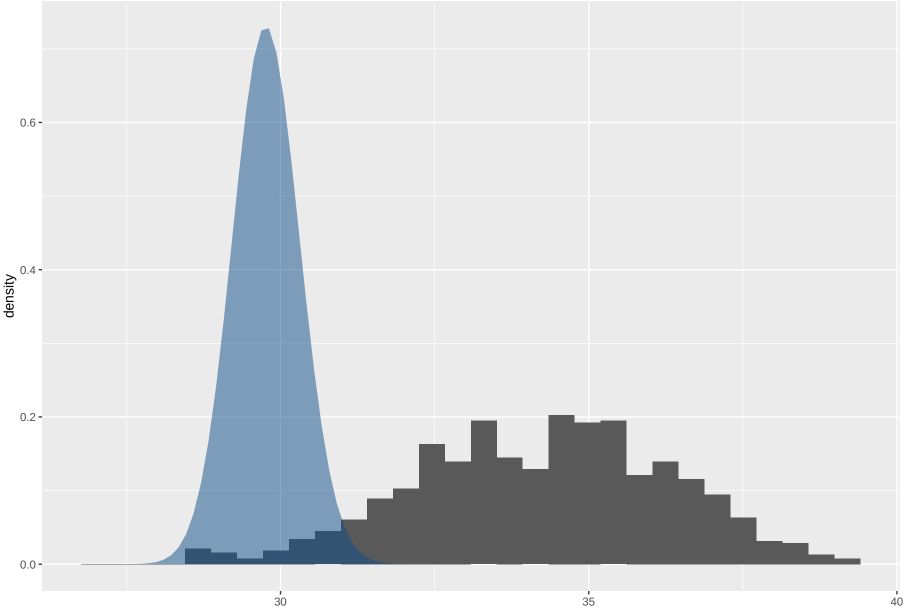
Bayesian Synthetic Likelihood, Further Exploration





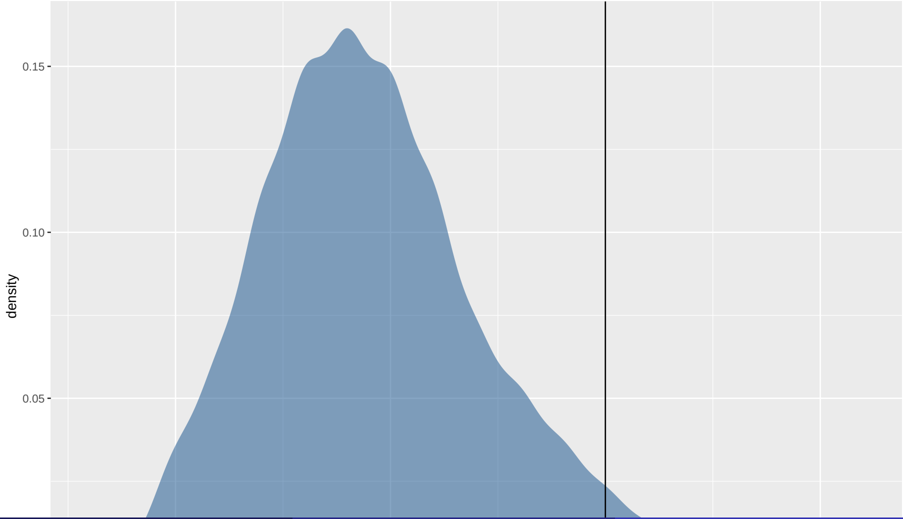


Bayesian Synthetic Likelihood, Further Exploration



Bayesian Synthetic Likelihood, Further Exploration

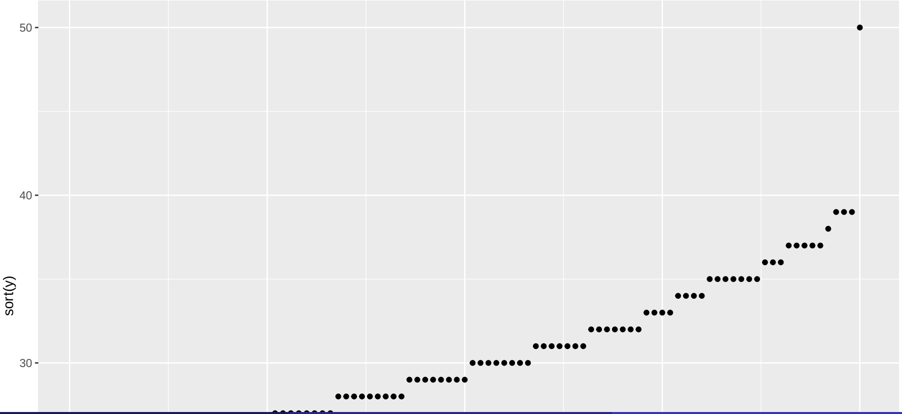
The synthetic likelihood approximation overestimated the true lambda. This is because the data had an unusually large maximal value.



Bayesian Synthetic Likelihood, Further Exploration

We attempted this on the toy Poisson model by using the coefficients of a polynomial regression on the ordered data as our statistics.

$$Y_{(i)} = \beta_0 + \beta_1 i + \beta_2 i^2 + \beta_3 i^3$$



example with ricker to show that it still works even when our summary statistics aren't exactly multivariate normal

comparison to ABC (or maybe this could go earlier?)

Synthetic Likelihood, Wood (2010), we can get rid of this or move it

- Statistical Problem: Fitting a model to data from chaotic biological systems using maximum likelihood can yield estimates that may be too sensitive to random noise
- Model Choices: This paper is inherently frequentist in nature and uses a parametric Ricker model of population growth over time as an example
- Computational Tools: Wood proposes a multivariate normal approximation to a vector of summary statistics and demonstrates its efficacy using Markov Chain Monte Carlo
- Other Approaches Available: Standard methodology up to that point was to use likelihood based approaches for statistical inference on model parameters but these methods were known to perform poorly on chaotic systems

Ricker Model

$$N_{t+1} = rN_t e^{-N_t + e_t} \qquad \text{where } e_t \sim N(0, \sigma_e^2)$$

$$Y_t | N_t \sim Poisson(\lambda = \phi N_t)$$

N_t is an autoregressive time-series modeling true, unknown population density at discrete time points. r is a population growth rate parameter, e_t is random noise. Y_t are the number of individuals sampled from the total population, ϕ is a scaling parameter.

Ricker Model

IMAGE

Ricker Model, our results

Paper pros and cons