Bayesian Statistics in Astrophysics

Jack Walton

October 23, 2019

Disclaimer

I am *not* talking about my own research because:

1. *Most* of you saw me talk at the applied PGR conference at the end of June
2. I have no new results since June (I suspended studies for 3 months over summer)

Figure 1: No new research here...

Table of contents

Bayesian & frequentist statistics

Frequentist statistics

- This approach to statistics will be familiar to most
- Think p-values, hypothesis testing, confidence intervals etc.
- However, it is not the only statistical framework (nor is it the focus of this talk…)

Bayesian vs. frequentist statistics

The difference between Bayesians and frequentists lies in the interpretation of probability…

For a *frequentist*:

> An event's probability is the limit of its relative frequency in many trials

For a *Bayesian*:

> An event's probability is a degree of belief

Why Bayesian?

- Philosophically aligns with how we practice science: *updating* our *belief*s in light of *new evidence*
- Allows the inclusion of expert information through a *prior distribution*
- For events that only occur once, how appropriate is a methodology which relies on repeatability?

Bayes' Theorem

$$\pi(\theta \mid \boldsymbol{x}) = \frac{\pi(\theta)\, L(\boldsymbol{x} \mid \theta)}{\int_\Theta \pi(\boldsymbol{x} \mid \theta)\, \mathrm{d}\theta}$$

- $\pi(\theta)$ represents our prior beliefs
- $L(\boldsymbol{x} \mid \theta)$ is the likelihood of observing $\boldsymbol{x}$ given the model & parameters θ
- $\int_\Theta \pi(\boldsymbol{x} \mid \theta)\, \mathrm{d}\theta$ is the normalising constant (probability of $\boldsymbol{x}$)
- $\pi(\theta \mid \boldsymbol{x})$ represents our posterior beliefs

Figure 2: Purportedly Bayes

Bayes' Theorem

Typically, $\int_\Theta \pi(\boldsymbol{x} \mid \theta)\, \mathrm{d}\theta$ is *very* difficult to compute.

Instead we often consider:

$$\pi(\theta \mid \boldsymbol{x}) \propto \pi(\theta) \times L(\boldsymbol{x} \mid \theta)$$
$$\text{posterior} \propto \text{prior} \times \text{likelihood}$$

Figure 2: Purportedly Bayes

MCMC

- MCMC — Markov Chain Monte Carlo
- Class of algorithms used to sample from probability densities
- We can use them to sample from $\pi(\theta \mid \boldsymbol{x})$, our posterior distribution
- Avoids the computation of $\pi(\boldsymbol{x})$

Stan

- Probabilistic programming language wrote in C++. Accessed via interfaces with Python, R, Matlab, Julia...
- Stan implements current state-of-the-art MCMC algorithms
- Named after Stanislaw Ulam, a mathematician and nuclear physicist and pioneer of Monte-Carlo methods.

Figure 3: Stanislaw & the FERMIAC

Sunspot occurrence: a case study

What *are* sunspots?

- Dark regions which appear on the surface of the sun
- Cooler areas, caused by concentrations of magnetic field flux
- Precursor to more dramatic events such as solar flares and coronal mass ejections
- Significant concern for astronauts living in space, airline passengers on polar routes and satellite engineers

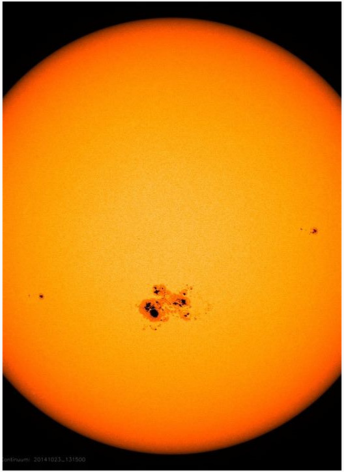


Figure 4: Sunspots

Data

We shall use the mean annual data for the International Sunspot number, under the responsibility of the Royal Observatory in Belgium since 1980.

Figure 5: Royal observatory of Belgium

Normal $AR(1)$ model

$$S_t \sim Normal(\mu_t, \sigma^2)$$
$$\mu_t = \varphi_1 + \varphi_2 S_{t-1}$$

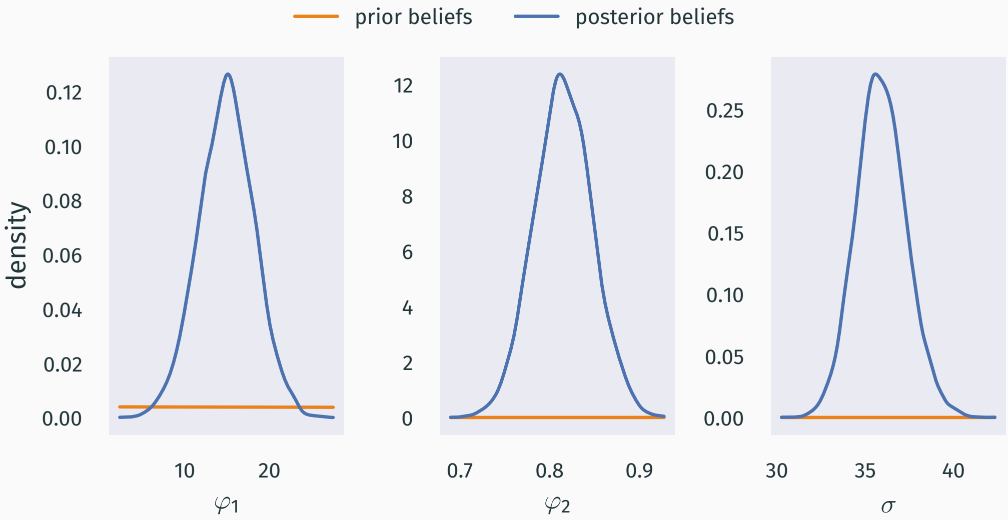
Given the observed data can we infer the parameters φ_1, φ_2 and σ?

Results: summary

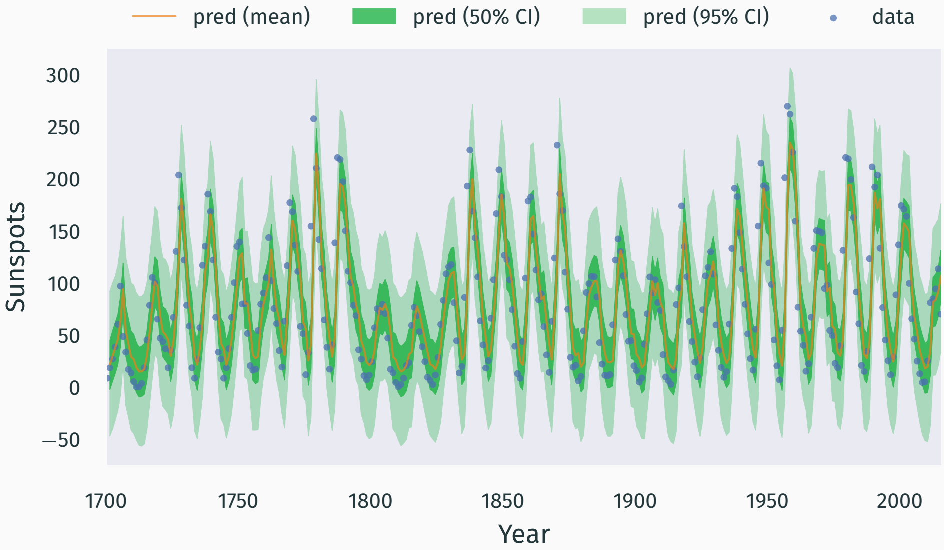
Parameter	mean	2.5%	97.5%	ESS
φ_1	14.90	8.35	21.33	4500
φ_2	0.82	0.75	0.88	4500
σ	35.91	33.20	38.81	5500

Table 1: Summary of posterior samples after running Stan for 10 000 iterations (3 seconds).

Results: posterior densities



Results: posterior predictives



Negative Binomial $AR(1)$ model

$$
\begin{aligned}
S_t &\sim NB(p_t, \theta) \\
p_t &= \theta / (\theta + \mu_t) \\
\log(\mu_t) &= \varphi_1 + \varphi_2 S_{t-1}
\end{aligned}
$$

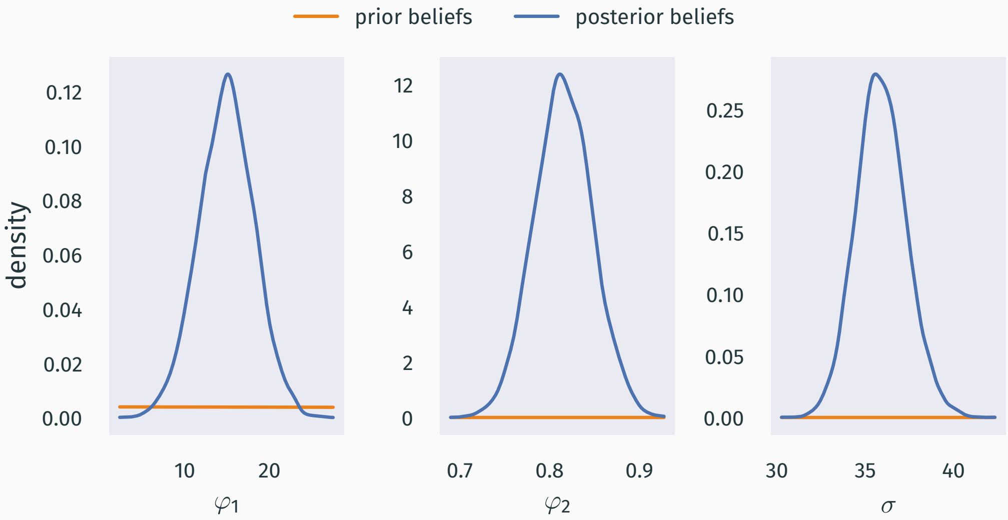
Given the observed data can we infer the parameters φ_1, φ_2 and θ?

Results: summary

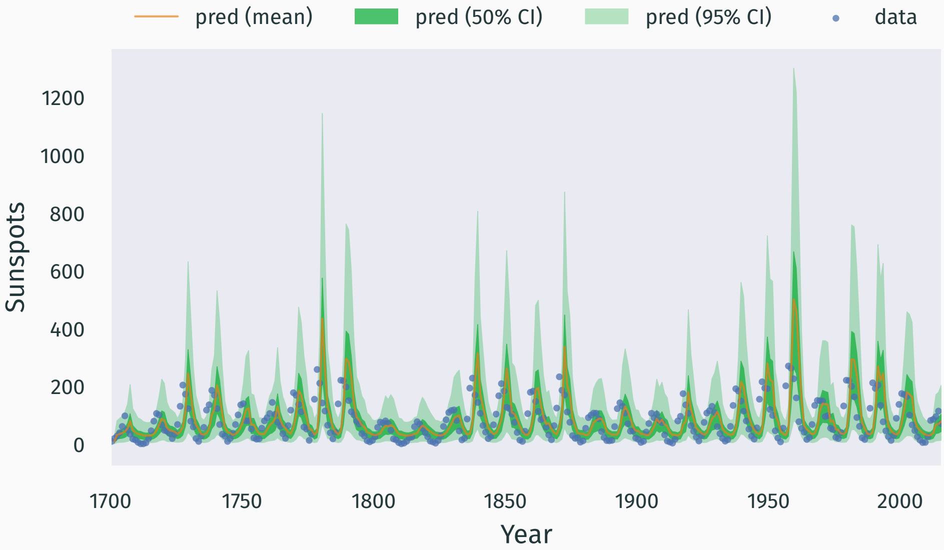
Parameter	mean	2.5%	97.5%	ESS
φ_1	3.33	3.21	3.46	5300
φ_2	0.01	0.01	0.01	6300
θ	2.55	2.16	2.99	5200

Table 2: Summary of posterior samples after running Stan for 10 000 iterations (30 seconds).

Results: posterior densities



Results: posterior predictives



Conclusion

- Modern computing power is making Bayesian methodologies more accessible
- Many ‘black-box’ MCMC implementations make inference pain-free
- The inclusion of prior information can be useful for astronomical events which have limited observational data

References

Joseph M Hilbe, Rafael S De Souza, and Emille EO Ishida.
Bayesian models for astrophysical data: using R, JAGS, Python, and Stan.
Cambridge University Press, 2017.

Thanks