Bayesian Statistics in Astrophysics

Jack Walton

November 26, 2019

Newcastle University

This talk's not *really* about astrophysics. . .

The second part of this talk makes a case study using astrophysical data

However, the methodolgy we'll explore can be used to model *any* observation which varies through time.

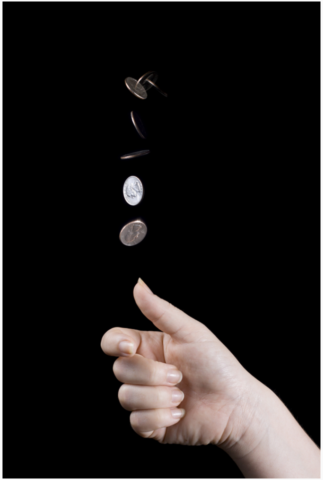
A tale of two parts

- Bayesian & frequentist statistics
    - The two approaches and how they differ
    - An introduction to MCMC and Stan
- Sunspot occurrence
    - What they are and why we should care
    - Autoregressive models
    - Model fitting and results

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 beliefs 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 1: 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 1: 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 2: Stanislaw & the FERMIAC

Sunspot occurrence: a case study

What *are* sunspots and who cares anyway?

- 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

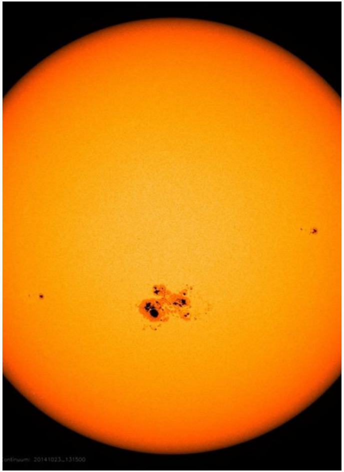


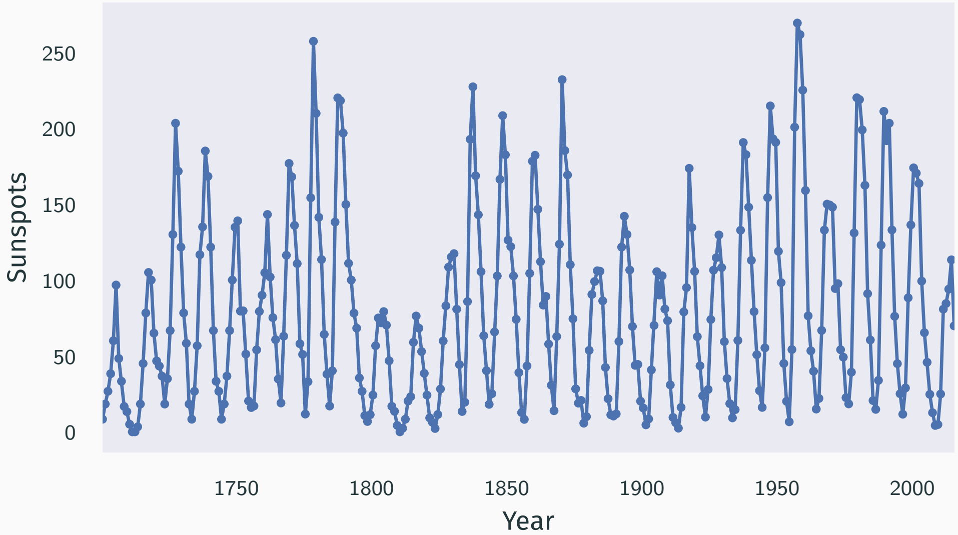
Figure 3: Sunspots

The data

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

Figure 4: Royal observatory of Belgium

The data



Sunspots
250
200
150
100
50
0
1750
1800
1850
1900
1950
2000
Year

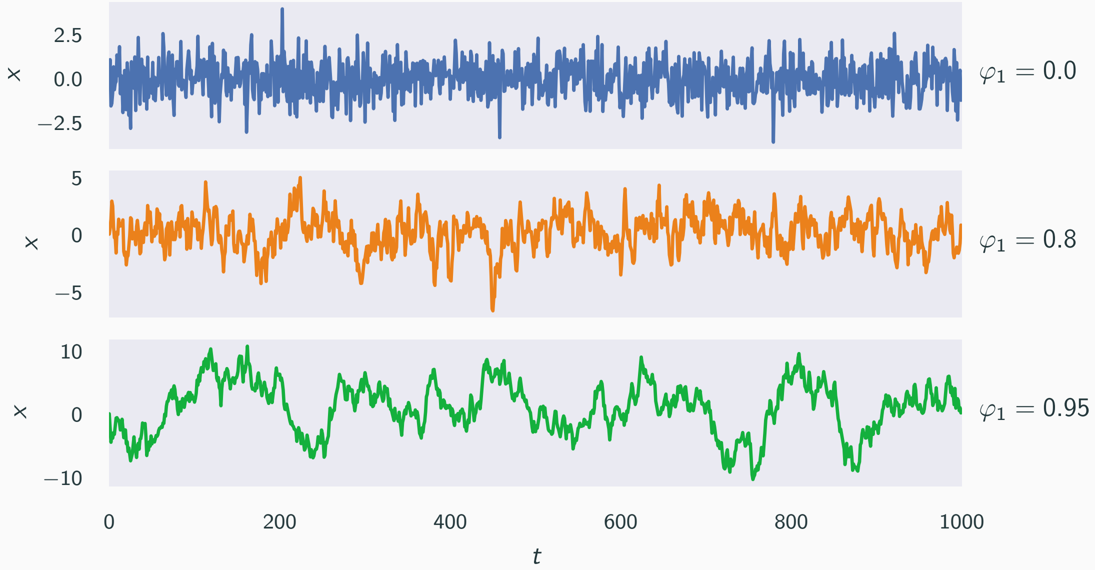
Autoregressive models

Autoregressive models predict future behaviour given past behaviour

An $AR(p)$ model:

$$X_t \sim Normal(\mu_t, \sigma^2)$$
$$\mu_t = \alpha + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \varphi_3 X_{t-3} + \ldots + \varphi_{t-p} X_{t-p}$$

In the flesh: $AR(1)$ processes



Normal $AR(1)$ model

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

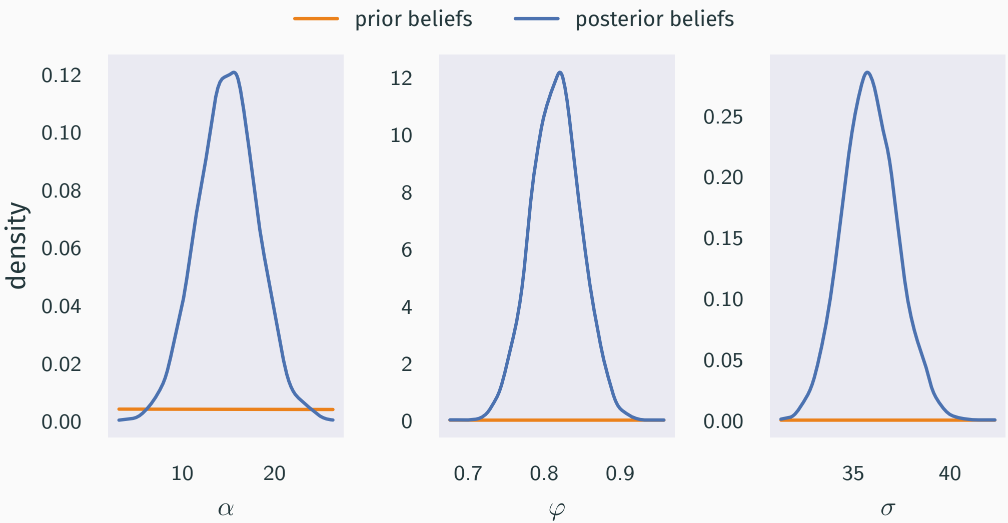
Given the observed data can we infer the parameters α, φ and σ?

Results: summary

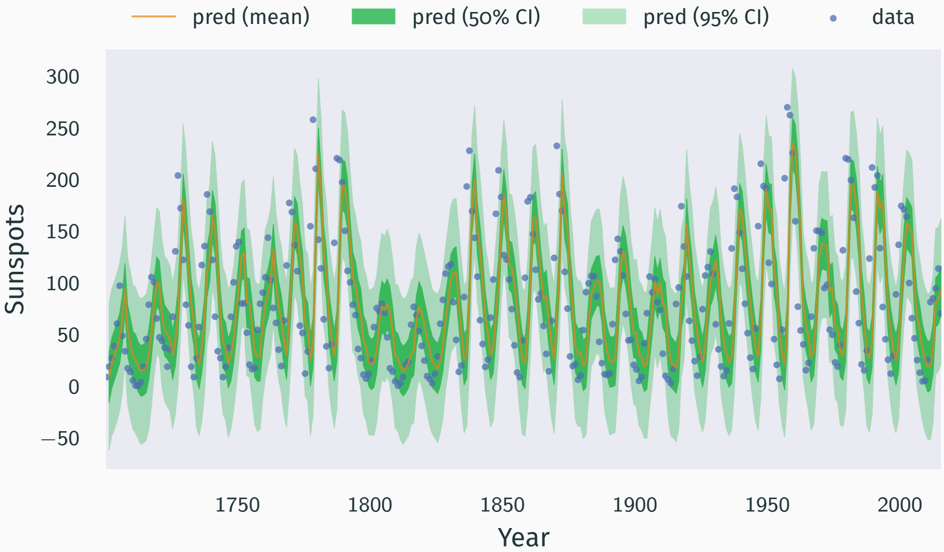
Parameter	mean	2.5%	97.5%	ESS
α	14.89	8.54	21.30	4800
φ	0.82	0.75	0.88	4900
σ	35.86	33.17	38.72	6400

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

Results: posterior densities



Results: posterior predictives

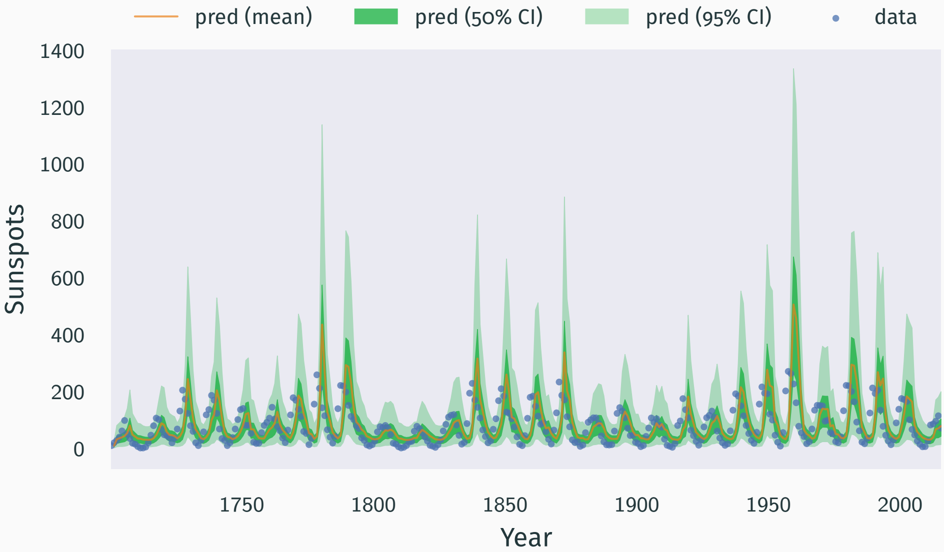


Negative Binomial $AR(1)$ model

$$S_t \sim NB(p_t, \theta)$$
$$p_t = \theta/(\theta + \mu_t)$$
$$\log(\mu_{t+1}) = \alpha + \varphi S_{t-1}$$

Given the observed data can we infer the parameters α, φ and θ?

Results: posterior predictives

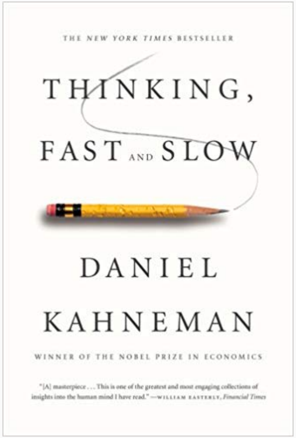


Conclusion

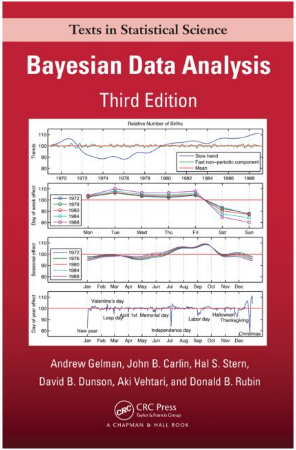
- Modern computing power is making Bayesian methodologies more accessible
- Many ‘black-box’ MCMC implementations make inference (relatively) pain-free
- The inclusion of prior information can be useful for events which have limited observational data

People who liked this also liked…

Suitable bed time reading:



Not suitable bed time reading:



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