Searching for signals in noise

Gregory Ashton, supervised by D.I. Jones & R. Prix

G.Ashton@soton.ac.uk

University of Southampton

I: Motivations

Astrophysics contains a huge number of interesting observed phenomena, let alone those that we have yet to see. The problem we often have is that that our observations are in a low signal to noise regime; this can mean several models can explain the observed data.

In this poster we discuss a method for quantitatively assessing how well several models fit some data. The aim being to decide, eventually, given the observed data which astrophysical model do we think is most likely.

II: Bayesian Data Analysis

While many methods exist for comparing some models, an intuitive approach can be found by using Bayes rule

$$p(\text{model}|\text{data}) = p(\text{data}|\text{model})\frac{p(\text{model})}{p(\text{data})} \tag{1}$$

where by $P(\text{model}|\text{data})$ we mean ”The probability of model, given that we have observed some data.

III: Model comparison

The issue with Bayes rule as written in eqn.(1) is that $p(\text{data})$ is often difficult, or even impossible to define. Instead, we can compare two models, say M_A and M_B by looking at the ratio

$$\frac{p(M_A|\text{data})}{p(M_B|\text{data})} = \frac{p(\text{data}|M_B)\,p(M_A)}{p(\text{data}|M_A)\,p(M_B)} \tag{2}$$

This ratio can directly be interpreted as the ’odds ratio’, or how much more we should believe model A of B given the data. The last fraction reflects our prior knowledge about the two models. Unless we have a strong reason to believe otherwise this is generally set to unity: $p(M_A) = p(M_B)$.

IV: Signals in noise

The key to data analysis is to define our ’likelihood’ function: the probability of the data given the model $p(\text{data}|\text{model})$. To calculate this we must first assume that the observed data is the sum of a a deterministic signal and a central Gaussian noise process. That is we can write the observed data as

$$W(t) = f(t;\vec{\theta}) + n(t;\sigma). \tag{3}$$

Here $f(t;\vec{\theta})$ is the signal model with parameters $\vec{\theta}$, while $n(t;\sigma)$ is the noise process with strength σ. If we subtract the signal model (with the correct parameters) from the observed data we should be left with Gaussian noise: $W(t) - f(t;\vec{\theta}) = n(t;\sigma)$. Then the probability of a single observed data point at t_i

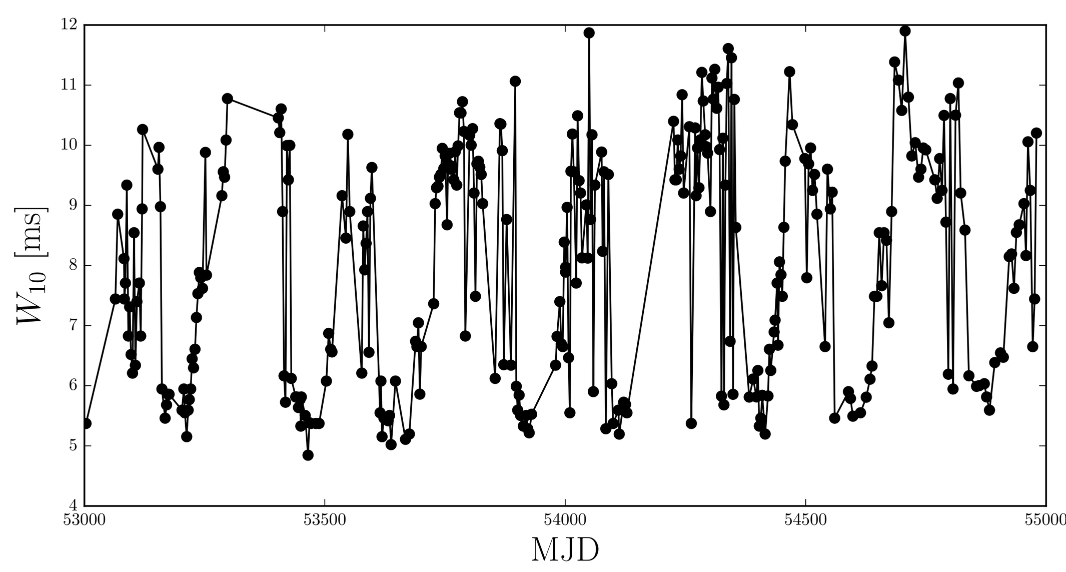
$$p(W(t_i)|\text{model}, \vec{\theta}, \sigma) = \frac{1}{\sqrt{2\pi\sigma}}e^{-\frac{(W(t_i)-f(t_i))}{2\sigma^2}} \tag{4}$$

Here the model, and hence the function $f(t)$, are yet to be determined. Given a sequence of N observations (e.g. some data) we can compute the likelihood function as

$$p(\text{data}|\text{model}, \vec{\theta}, \sigma) = \prod_{i=1}^{N} p(W(t_i)|\text{model}), \vec{\theta}, \sigma \tag{5}$$

V: Example

To illustrate how we can apply Bayesian data analysis consider the data shown in the figure below. This is a plot of the measured beam width of pulsar B1828-11 showing a distinct periodic behaviour. This data was originally published in figure 5 of [1], we are thankful to the original authors for allowing us access to this data.



Prior to this paper, slow periodic modulation of the spin-down was cited as evidence for precession in this pulsar. Clearly the beam width is also being modulated at this same time-scale. The authors argue that the beam width is not smoothly varying between, but instead *switching*. This has important implications for neutron star physics.

[1] Lyne, A., Hobbs, G., Kramer, M., Stairs, I., and Stappers, B. (2010). *Switched Magnetospheric Regulation of Pulsar Spin-Down.* Science, 329:408âĂŞ.

VI: Model A

The simplest model we can propose to explain the periodic features is a simple trig. function. This does not capture all of the physics of alternative models, but nevertheless will test the assumption that W_{10} switches instantaneously between two values. The signal function for this model is

$$f(t; W_0, A, f, \phi_0) = W_0 + A\sin(2\pi f t + \phi_0) \tag{6}$$

VII: Model B

The model proposed by the authors can be captured by a simple square wave. We then make an assumption that the observed data is the sum of a square wave and noise.

$$g(t; W_0, A, f, \phi_0) = W_0 + \text{square}(t\; W_0, A, f, \phi_0) \tag{7}$$

VIII: Marginalisation and odds ratio

Plugging the model functions 6 and 7 into the likelihood 5 we have $P(\text{data}|M_i, \vec{\theta}, \sigma)$. The next step is to marginalise over all the model parameters:

$$p(\text{data})|M_i) = \int\int p(\text{data}|M_i, \vec{\theta}, \sigma)p(\vec{\theta})p(\sigma)d\vec{\theta}d\sigma \tag{8}$$

Often this integral will be unfeasible analytically, instead we can turn to numerical method such as MCMC and nested sampling.

IX: Results

The observed beam width data was tested with both a sinusoidal model and a square-wave model. We computed the marginal likelihoods using MCMC software *dan.iel.fm/emcee*. This resulted in an odds ratio:

$$\log_{10}\left(\frac{p(M_A|\text{data})}{p(M_B|\text{data})}\right) = -6 \tag{9}$$

This establishes that the square wave model proposed by the authors significantly out performs a sinusoidal model. We are currently in the process of exploring more physical models to further probe the assumptions.