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 is full of interesting and observed phenomena. The problem we often have is that our observations are in a low signal to noise regime; this can mean it is difficult to choose between which of several models explains what we have observed.

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

II: Bayesian Data Analysis

We have some data, and several models $\{M_1, \ldots, M_n\}$ which could explain it. An intuitive approach to decide between them be found by writing Bayes rule for the i^{th} model given some data:

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

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

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

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

This ratio can directly be interpreted as the 'odds ratio', or how much more we should believe model A over 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

To compute the odds ratio between two models, we need to first define our 'likelihood' function $p(\text{data}|M_i)$: *the probability of the data given the model*. To calculate this, we first calculate the conditional probability distribution for all the parameters.

For signals in noise we assume that the observed data is the sum of a deterministic signal and a central Gaussian noise process:

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

Here $f(t;\vec{\theta})$ is the *signal function* 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 *will* 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, given parameters $\vec{\theta}$ and σ is:

$$p(W(t_i)|M_i,\vec{\theta},\sigma) = \frac{1}{\sqrt{2\pi\sigma}}\exp\left\{-\frac{(W(t_i) - f(t_i;\vec{\theta})}{2\sigma^2}\right\} \tag{4}$$

Here the model, and hence the function $f(t;\vec{\theta})$, are yet to be determined. Given a sequence of N observations (i.e. some data) we can compute the conditional probability distribution of the data:

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

V: Marginalisation

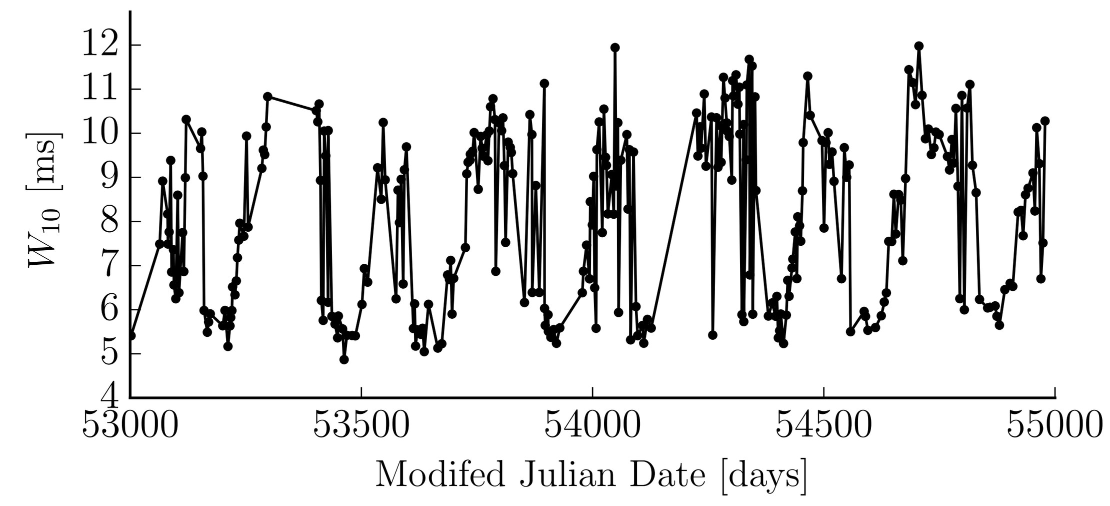
Plugging the model functions (such as (7) and (8)) into eqn. (5) we have the conditional likelihood. To compute the likelihood we marginalise over all the parameters:

$$p(\text{data})|M_i) = \iint p(\text{data}|M_i,\vec{\theta},\sigma)p(\vec{\theta})p(\sigma)d\vec{\theta}d\sigma \tag{6}$$

Often this integral will be unfeasible analytically, instead we can turn to numerical methods such as MCMC and nested sampling. We also need to specify our prior distributions on the parameters $p(\vec{\theta})$ and $p(\sigma)$, often a simple uniform prior is sufficient. Having computed this for two different models, they can be directly compared using eqn. (2).

VI: 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] and we are thankful to the authors for allowing us access to this data.



From this data, the authors argue that the beam width W_{10} is periodically *switching* between two distinct values. This is in stark contrast to the standard view that the beam width should be *smoothly* oscillating. The physics which underlies these two ideas are inherently different, so which of these models best explains the data is an important question 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.

VII: Model A

The simplest way to model smoothly varying periodic features is with a 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 given by:

$$f(t;W_0,A,f,\phi_0) = W_0 + A\sin(2\pi ft + \phi_0). \tag{7}$$

VIII: Model B

The model proposed by the authors can be captured by a simple square wave which we give here as a generic function:

$$g(t;W_0,A,f,\phi_0) = W_0 + A\,\text{square}(t\ f,\phi_0). \tag{8}$$

Where the parameters can be directly compared with those in model A. This model can be improved by allowing for a duty cycle as well.

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 from *dan.iel.fm/emcee*. This resulted in an odds ratio:

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

This establishes that the square wave model proposed by the authors significantly outperforms a sinusoidal model.

In the figure to the right, several realisations of the two models are displayed along with the original data. This can be a useful way to evaluate the fit of the models.

The next step is to add more physics into the signal models. This is to be done by predicting the beam width variations due to realistic modulations such as free precession and a biased magnetosphere.