TSDA

- Change point algorithms - non-periodic time series
- Gaussian process modelling (optional)

- Both these use a “Bayesian approach”

Bayesian approaches

- **Initial beliefs** — what do we know already?
- **Objective data** — do the experiment
- This leads to: **New and improved belief**
- This is the initial belief **the next time**
- Builds on our existing knowledge — new information ++ (quantifiable)

- This is just what we do in real life
  Intuitive, philosophical. No p-values, student t-tests etc

Based on Jon Zwart's talk, NASSP 2017

Bayes's theorem

$$\text{prob}(B \mid A) = \frac{\text{prob}(A \mid B)\text{prob}(B)}{\text{prob}(A)}$$

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}}$$

$$P(\Theta \mid D) = \frac{L(D \mid \Theta)\Pi(\Theta)}{Z(\Theta)}$$

D = *known*, data

Θ= *model* parameters, eg. period

Π = *prior*, assumption

Based on Jon Zwart's talk, NASSP 2017

Bayes's theorem - example 1

A patient goes to see a doctor. The doctor performs a test with 98% reliability—that is, 98% of people who are sick test positive and 98% of the healthy people test negative. The doctor knows that only 5% of the people in the country are sick (prevalence of the disease).

If the patient tests positive, what are the chances the patient is sick?

Let's explore this using numbers, a population of 10000 people.

The numbers in the table are calculated from the information above.

Tests	Disease	Healthy		
+	490	190	680	**Total # of positive tests in the population**
-	10	9310	9320	**Total # of negative tests in the population**
	500	9500	10000	
	Total # of diseased in population	**Total # of healthy in population**		

Bayes's theorem - example 1

If the patient tests positive, what are the chances the patient is sick?

$$\text{prob}(A\,|\,B) = \frac{\text{prob}(B\,|\,A)\text{prob}(A)}{\text{prob}(B)}$$

$$p(\text{disease}\,|\,+) = \frac{p(+\,|\,\text{disease})p(\text{disease})}{p(+)}$$

$$= \frac{0.98 \times 0.05}{0.068} = 0.72$$

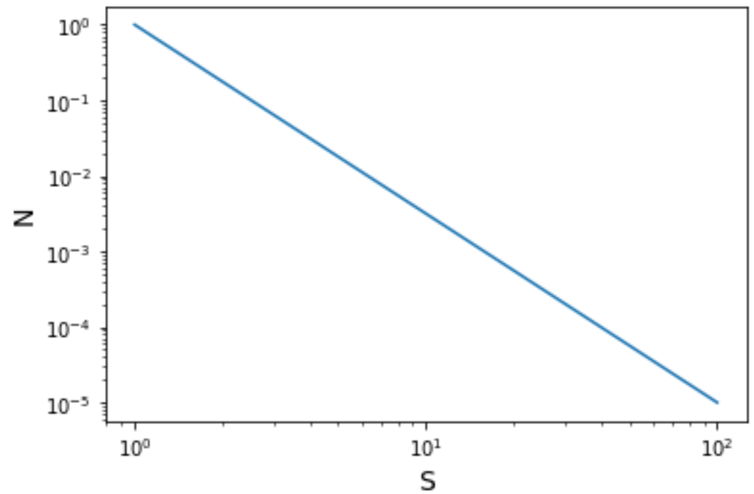
From the table on previous slide:
p(+|disease) = 0.98
p(disease) = 0.05
p(+) = (490 + 190)/10000 = 0.068

If the patient tests positive, there is a 72% chance that they are sick, given the efficacy of the test and the prevalence of the disease.

Bayes's theorem - example 2

We observe the sky with a radio telescope. Our model of the data (an event labelled *D*, consisting of a single measured flux density *f*) is that it is distributed in a Gaussian way about the true flux density *S* with a variance σ^2. The literature tells us the a-priori distribution of *S*, which we approximate here by the relation

$$\text{prob}(S) = KS^{-5/2}$$



Many more faint sources than bright ones

K just normalises the counts to unity; we presume one source in the beam at some flux-density level.

The probability of observing *f* when the true value is *S* is (i.e. prob(*D*|*S*)):

$$\exp\left[-\frac{1}{2\sigma^2}(f - S)^2\right]$$

Bayes's theorem - example 2

Using Bayes's theorem we calculate the probability of *S* given *D*:

$$\mathrm{prob}(S\,|\,D) = \frac{\mathrm{prob}(D\,|\,S)\mathrm{prob}(S)}{\mathrm{prob}(D)}$$

$$\mathrm{prob}(S\,|\,D) = K' \exp\left[-\frac{1}{2\sigma^2}(f - S)^2\right] S^{-5/2}$$

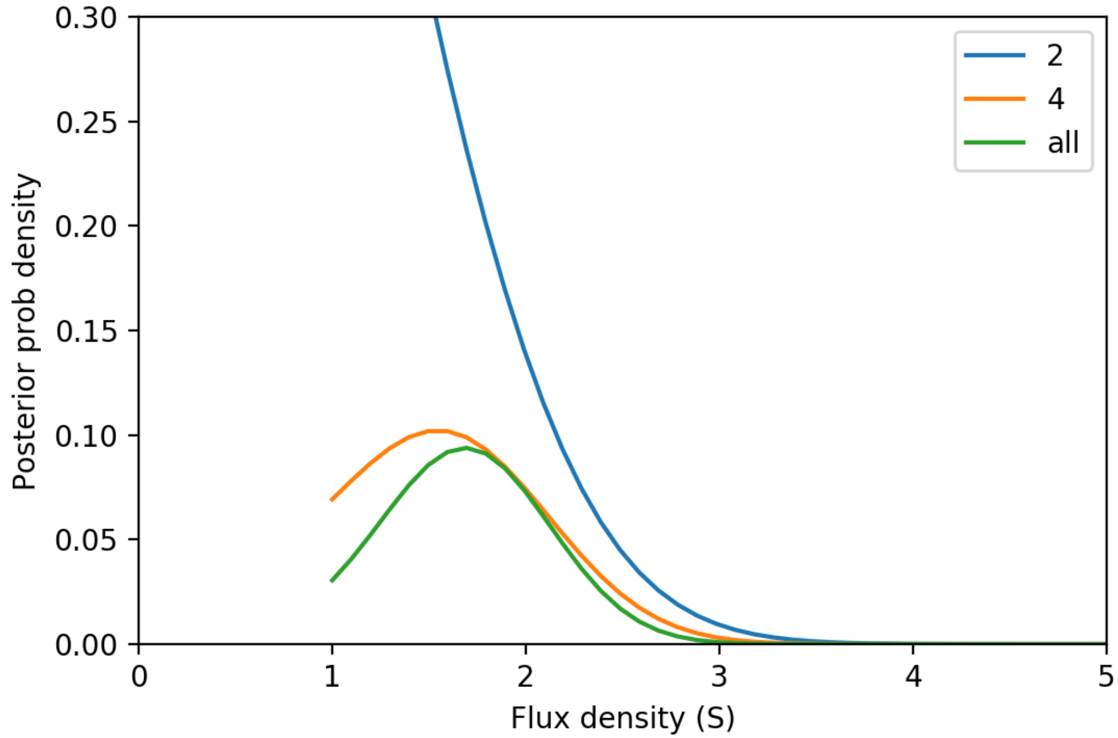
If our data, *D*, comprised *n* independent flux measurements, f_i then

$$\mathrm{prob}(S\,|\,D) = K' \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(f - S)^2\right] S^{-5/2}$$

Bayes's theorem - example 2

Let's use a specific example where source counts range from 1 to 100 units, noise is $\sigma = 1$, and the data (i.e. f) were 2, 1.3, 3, 1.5, 2, 1.8.

Calculate the posterior probabilities for a range of flux densities, S.



For fewer measurements, the probability is strongly affected by the prior (prob(S) = $S^{-5/2}$) but as we add more data, the probability moves towards a Gaussian around S~1.8.