3. Parameter Estimation and Goodness of Fit – Part Two

In the previous section we have discussed how to estimate parameters of an underlying pdf model from sample data.

We now consider the closely related question:

How good is our pdf model in the first place?

3. Parameter Estimation and Goodness of Fit – Part Two

In the previous section we have discussed how to estimate parameters of an underlying pdf model from sample data.

We now consider the closely related question:

How good is our pdf model in the first place?

Simple Hypothesis test – example.

Null hypothesis:	sampled data are drawn from a normal pdf, with mean μ_{model} and variance σ^2.

We want to **test** this null hypothesis: are our data consistent with it?

Assume (for the moment) that σ^2 is known.

Example

Measured data: $\{x_i : i = 1,\ldots,10\}$ $\sum_{i=1}^{10} x_i = 47.8$

Null hypothesis: $x \sim N(\mu, \sigma^2)$ with $\mu_{\text{model}} = 4$

Assume: $\sigma = 2$ $\sigma_\mu^{\ 2} = 0.4$

Under NH, sample mean $\boxed{\bar{x}_{\text{model}} \sim N(4, 2^2/10)}$

Observed sample mean $\bar{x}_{\text{obs}} = 4.78$

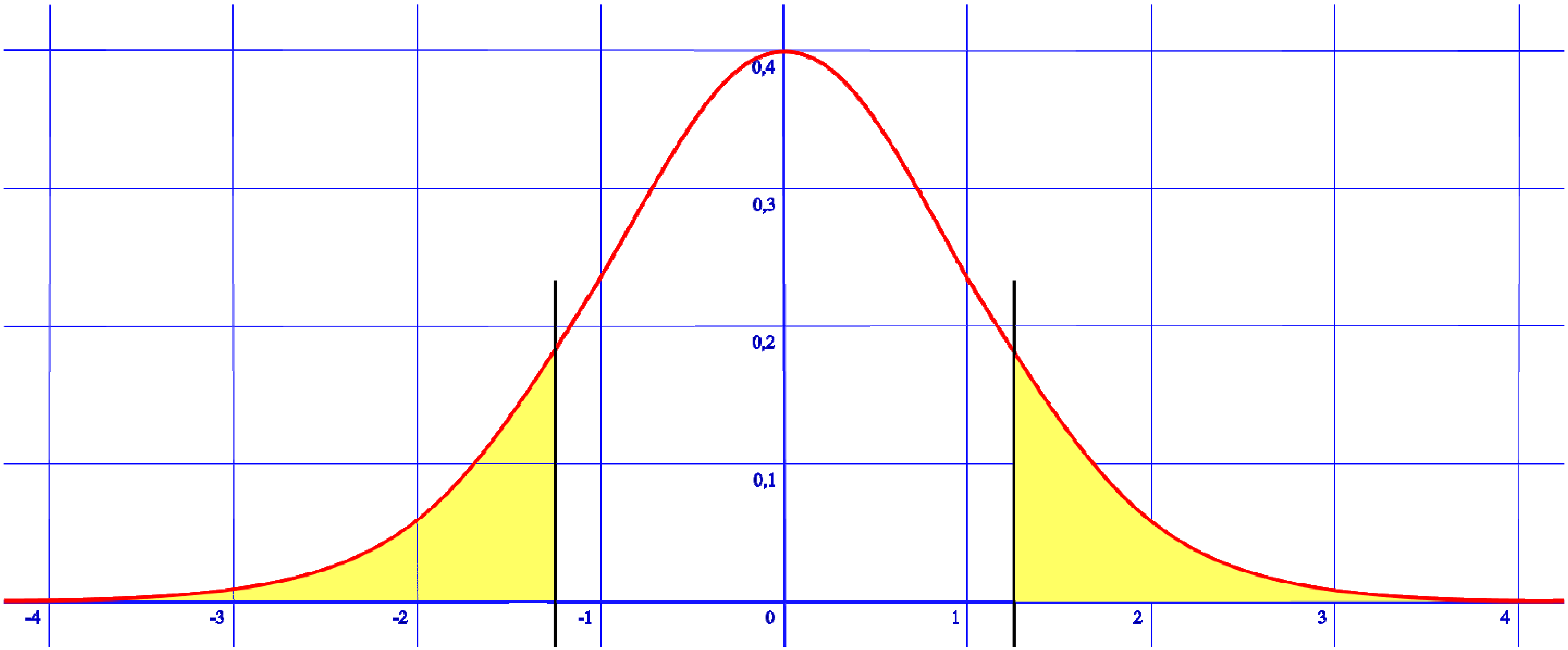
We transform to a **standard normal variable**

Under NH: $$Z = \left(\frac{\bar{x}_{\text{obs}} - \bar{x}_{\text{model}}}{\sigma_{\mu}} \right) \sim N(0,1)$$

From our measured data: $$Z_{\text{obs}} = \frac{4.78 - 4}{\sqrt{0.4}} = 1.233$$

If NH is true, how probable is it that we would obtain a value of Z_{obs} as large as this, or larger?

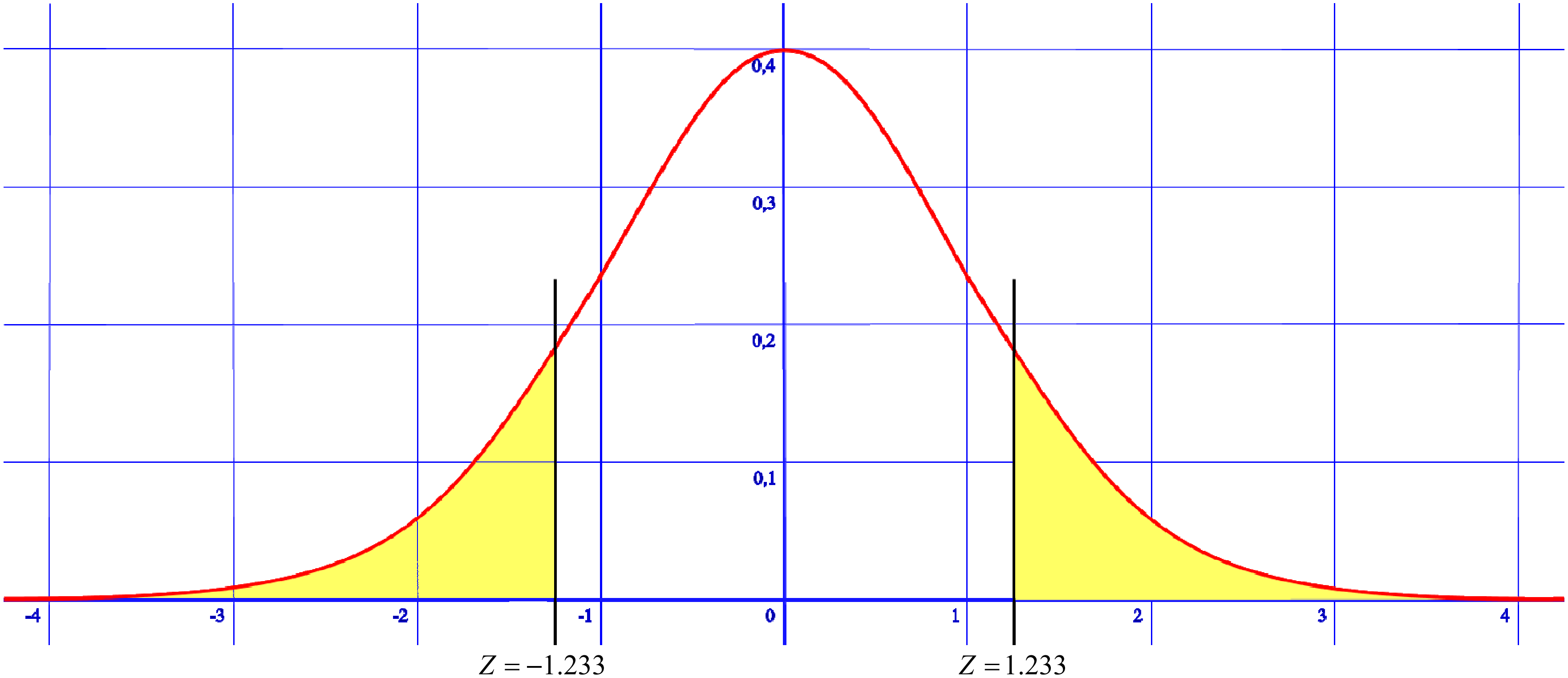
We call this probability the p-value



$$\text{p-value} \quad = \quad \mathrm{Prob}\left(|Z| \geq |Z_{\mathrm{obs}}|\right) \quad = \quad 1 - \int_{-Z_{\mathrm{obs}}}^{Z_{\mathrm{obs}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\tfrac{1}{2} z^2\right) dz$$

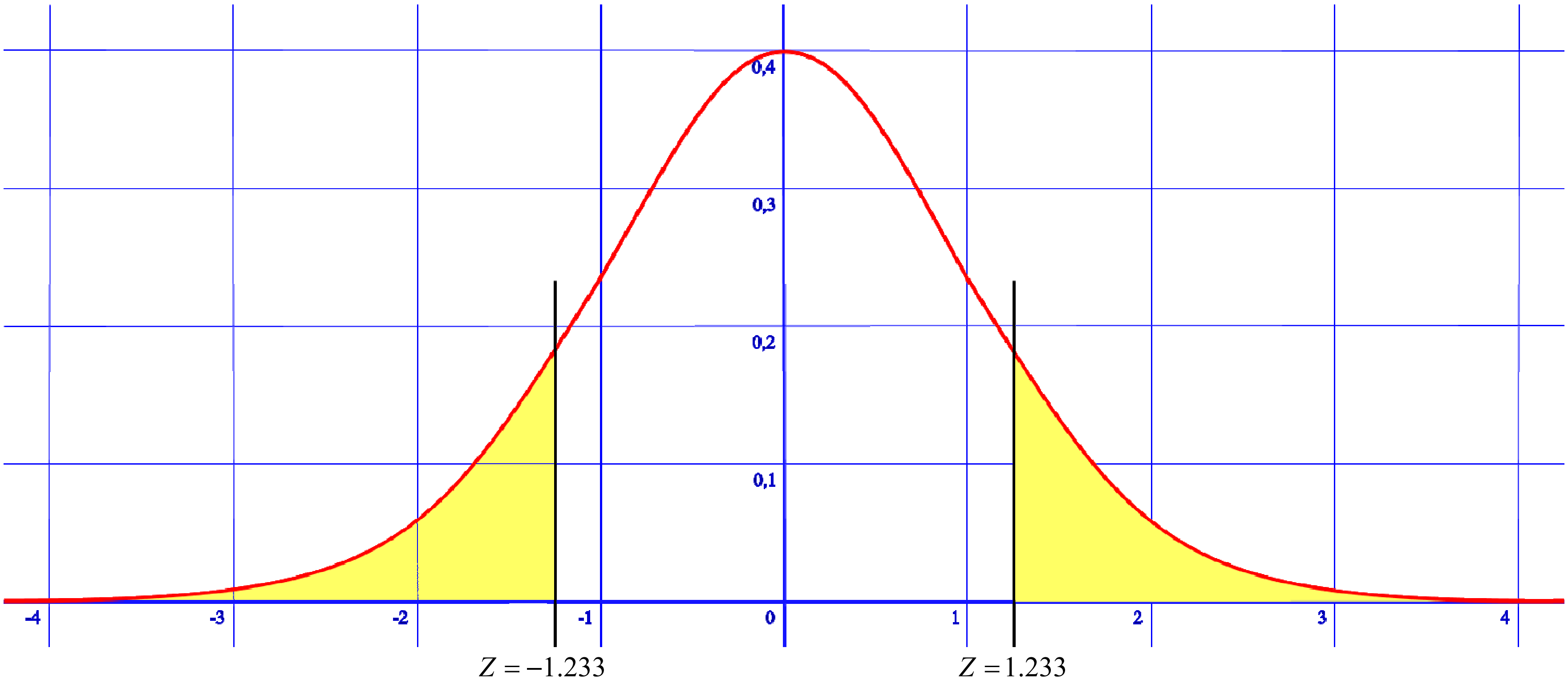
Simple programs to perform this probability integral (and many others) can be found in numerical recipes, or built into e.g. MATLAB or MAPLE.

Java applets also available online at http://statpages.org/pdfs.html



$$\text{p-value} = \mathrm{Prob}\left(|Z| \geq |Z_{\mathrm{obs}}|\right) = 0.2176$$

The ***smaller*** the p-value, the less credible is the null hypothesis.



$$\text{p-value} \quad = \quad \text{Prob}\left(|Z| \geq |Z_{\text{obs}}|\right) \quad = \quad 0.2176$$

The ***smaller*** the p-value, the less credible is the null hypothesis.

(We can also carry out a *one-tailed* hypothesis test, if appropriate, and for statistics with other sampling distributions).

Question 6: A one-tailed hypothesis test is carried out. Under the NH the test statistic has a uniform distribution $U[0,1]$.

The observed value of the test statistic is 0.8.

The p-value is:

A 0.8

B 0.9

C 0.2

D 0.1

Question 6: A one-tailed hypothesis test is carried out. Under the NH the test statistic has a uniform distribution $U[0,1]$.

The observed value of the test statistic is 0.8.

The p-value is:

A 0.8

B 0.9

$\boxed{\textbf{C} \quad 0.2}$

D 0.1

What if we *don't* assume that σ^2 is known?

We can estimate it from our observed data (provided $n \geq 2$)

We form the statistic $t_{\text{obs}} = \left(\dfrac{\bar{x}_{\text{obs}} - \bar{x}_{\text{model}}}{\hat{\sigma}_\mu} \right)$

where $$\hat{\sigma}_\mu{}^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} \left(x_i - \bar{x}_{\text{obs}} \right)^2$$

Accounts for the fact that we don't know μ , but must use $\bar{x}_{\text{obs}}$ when we estimate σ_μ

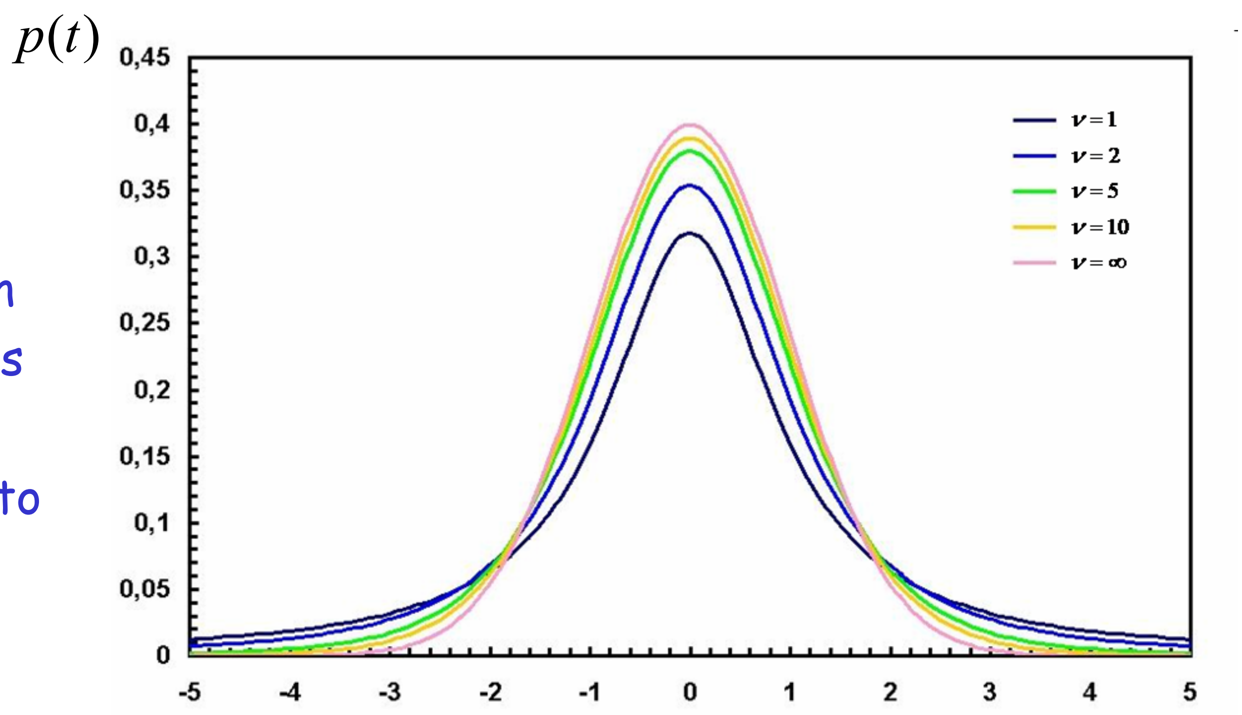
However, now t_{obs} no longer has a normal distribution.

In fact t_{obs} has a pdf known as the **Student's t distribution**

$$p(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$$

where $\nu = n-1$ is the no. degrees of freedom and $\Gamma(\nu) = \int_0^\infty x^{\nu-1}e^{-x}dx$

For small n the Student's t distribution has more extended tails than Z, but as $n \to \infty$ the distribution tends to $N(0,1)$



Question 7: The more extended tails of the students' t distribution mean that, under the null hypothesis

A larger values of the test statistic are more likely

B larger values of the test statistic are less likely

C smaller values of the test statistic are more likely

D smaller values of the test statistic are less likely

Question 7: The more extended tails of the students’ t distribution mean that, under the null hypothesis

A larger values of the test statistic are more likely

B larger values of the test statistic are less likely

C smaller values of the test statistic are more likely

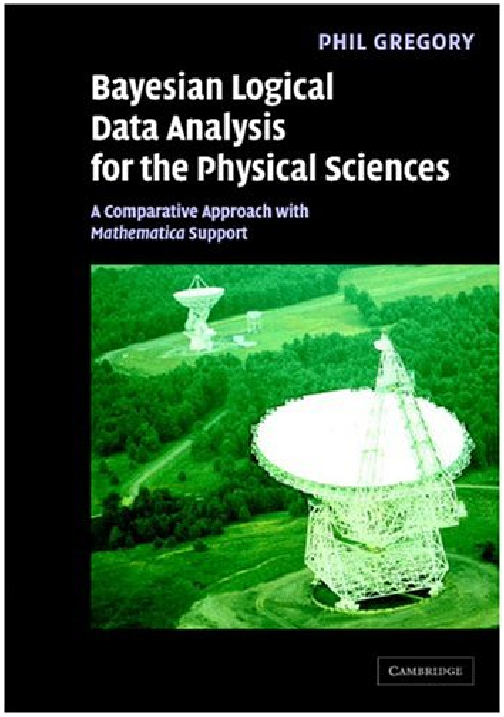
D smaller values of the test statistic are less likely

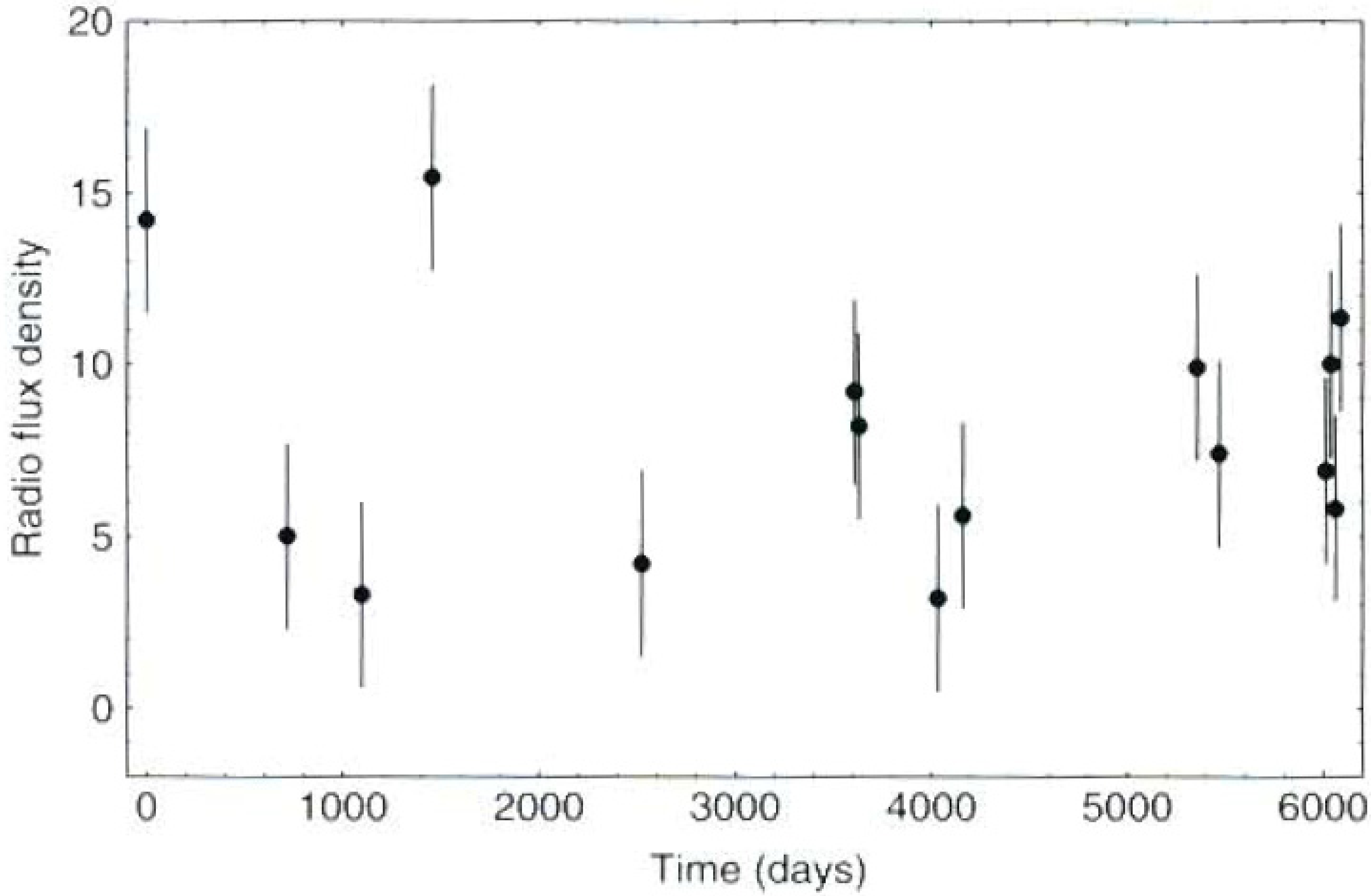
3. Parameter Estimation and Goodness of Fit – Part Two

More generally, we now illustrate the frequentist approach to the question of how good is the fit to our model, using the Chi-squared goodness of fit test.

We take an example from Gregory (Chapter 7)

(book focusses mainly on Bayesian probability, but is very good on frequentist approach too)





Model: radio emission from a galaxy is constant in time.

Assume residuals are iid, drawn from $N(0,\sigma)$

Goodness-of-fit Test: the basic ideas

1. Choose as our null hypothesis that the galaxy has an unknown but constant flux density. If we can demonstrate that this hypothesis is absurd at say the 95% confidence level, then this provides indirect evidence that the radio emission is variable. Previous experience with the measurement apparatus indicates that the measurement errors are independently normal with a $\sigma = 2.7$.
2. Select a suitable statistic that (a) can be computed from the measurements, and (b) has a predictable distribution. More precisely, (b) means that we can predict the distribution of values of the statistic that we would expect to obtain from an infinite number of repeats of the above set of radio measurements under identical conditions. We will refer to these as our hypothetical reference set. More specifically, we are predicting a probability distribution for this reference set.

   To refute the null hypothesis, we will need to show that scatter of the individual measurements about the mean is larger than would be expected from measurement errors alone.
3. Evaluate the χ^2 statistic from the measured data. Let's start with the expression for the χ^2 statistic for our data set:

$$\boxed{\chi^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{\sigma^2},}$$

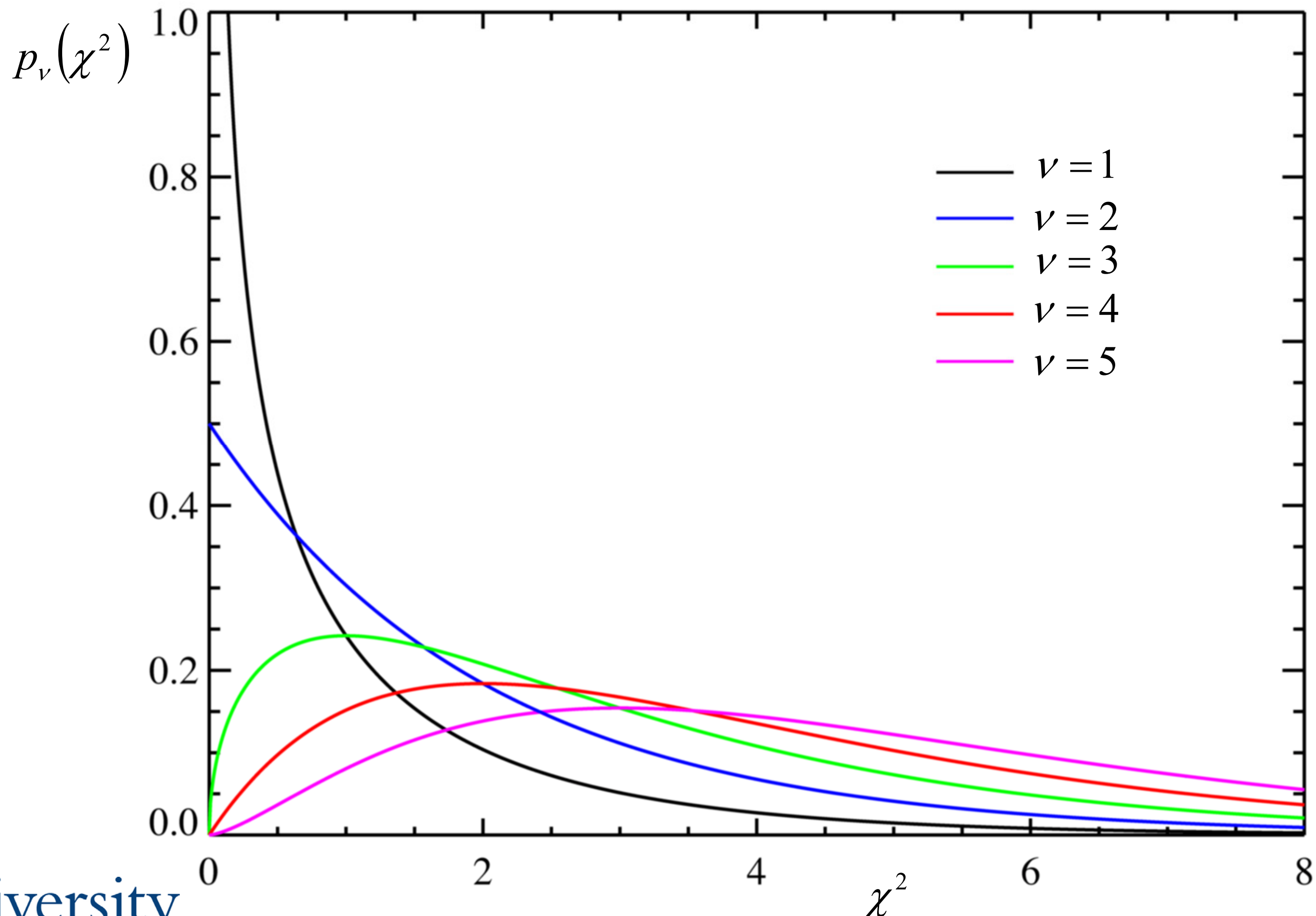
From Gregory, pg. 164

The χ^2 pdf

$$\boxed{p_\nu(\chi^2) = p_0 \times (\chi^2)^{\frac{\nu}{2}-1} e^{-\chi^2/2}}$$

Here ν is known as the number of **degrees of freedom** of the pdf.

The mean value of the pdf is ν and the variance is 2ν.

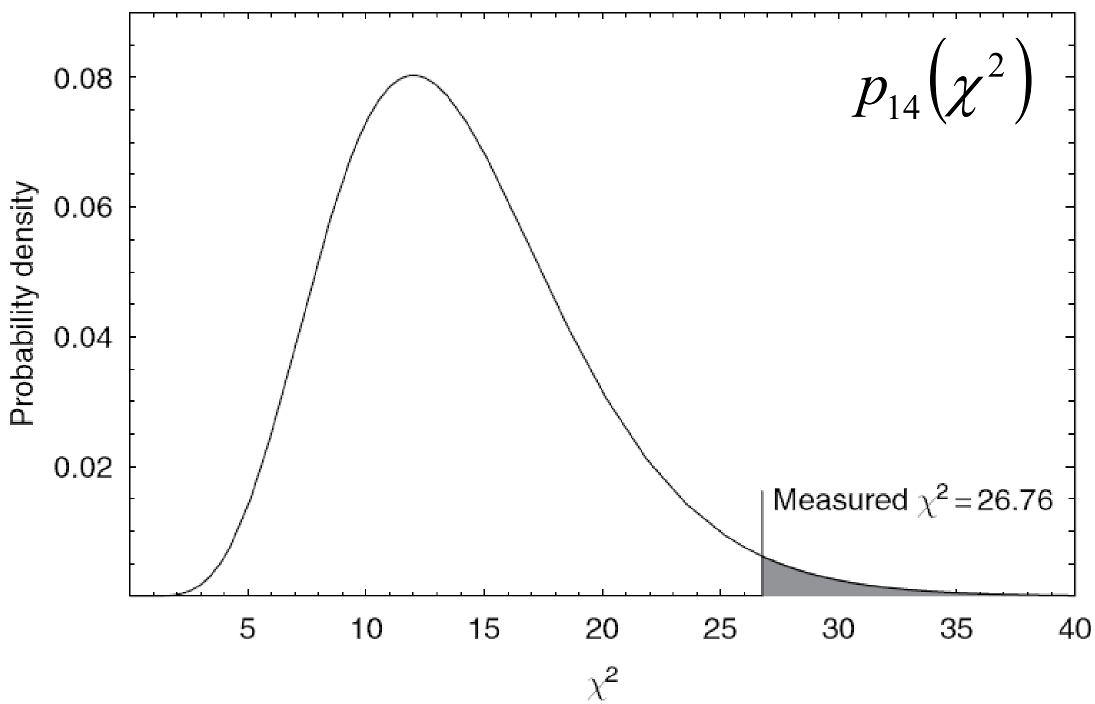


$$\chi^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(x_i - 7.98)^2}{2.7^2} = 26.76.$$

Day Number	Flux Density (mJy)
0.0	14.2
718.0	5.0
1097.0	3.3
1457.1	15.5
2524.1	4.2
3607.7	9.2
3630.1	8.2
4033.1	3.2
4161.3	5.6
5355.9	9.9
5469.1	7.4
6012.4	6.9
6038.3	10.0
6063.2	5.8
6089.3	11.4

$n = 15$ data points, but $\nu = 14$ degrees of freedom, because χ^2 statistic involves the *sample mean* and not the true mean.

We subtract one d.o.f. to account for this.

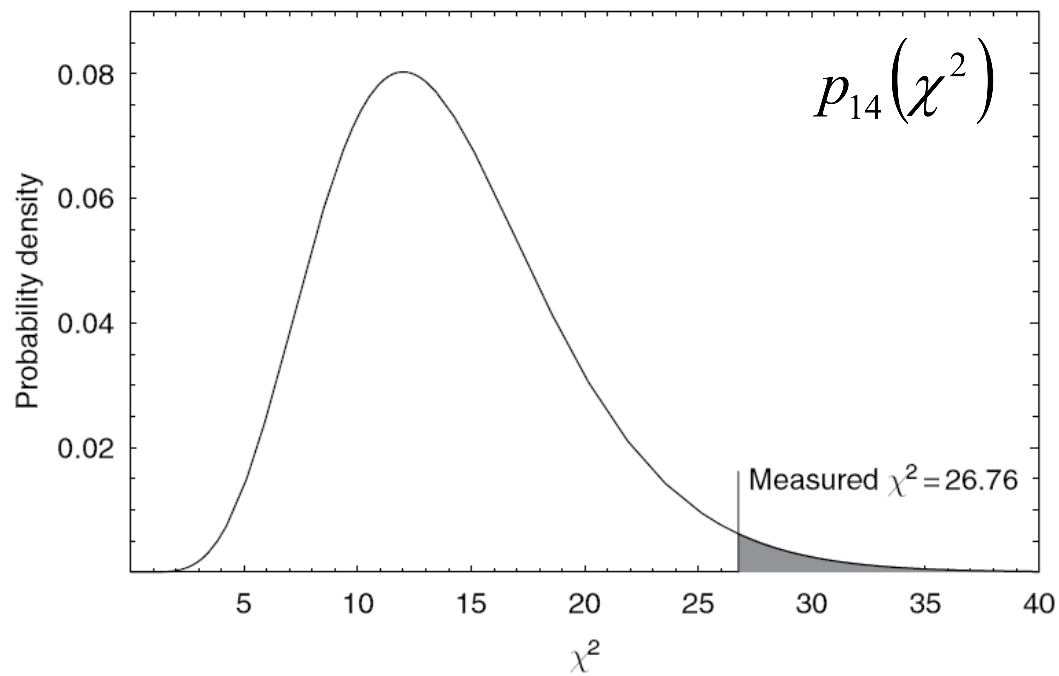


$$\chi^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(x_i - 7.98)^2}{2.7^2} = 26.76.$$

Day Number	Flux Density (mJy)
0.0	14.2
718.0	5.0
1097.0	3.3
1457.1	15.5
2524.1	4.2
3607.7	9.2
3630.1	8.2
4033.1	3.2
4161.3	5.6
5355.9	9.9
5469.1	7.4
6012.4	6.9
6038.3	10.0
6063.2	5.8
6089.3	11.4

$n = 15$ data points, but $\nu = 14$ degrees of freedom, because χ^2 statistic involves the *sample mean* and not the true mean.

We subtract one d.o.f. to account for this.



If the null hypothesis is true, how probable is it that we would measure as large, or larger, a value of χ^2?

If the null hypothesis were true, how probable is it that we would measure as large, or larger, a value of χ^2?

Recall that we refer to this important quantity as the **p-value**

$$\text{p-value} \;=\; 1 - P\left(\chi^2_{\text{obs}}\right) \;=\; 1 - \int_0^{\chi^2_{\text{obs}}} p_0 \, x^{\frac{\nu}{2}-1} \exp\left(-\frac{x}{2}\right) dx \;=\; 0.02$$

If the null hypothesis were true, how probable is it that we would measure as large, or larger, a value of χ^2?

Recall that we refer to this important quantity as the **p-value**

$$\text{p-value} \;=\; 1 - P\left(\chi^2_{\text{obs}}\right) \;=\; 1 - \int_0^{\chi^2_{\text{obs}}} p_0\, x^{\frac{\nu}{2}-1} \exp\left(-\frac{x}{2}\right) dx \;=\; 0.02$$

What precisely does the p-value mean?

> "If the galaxy flux density really *is* constant, and we repeatedly obtained sets of 15 measurements under the same conditions, then only 2% of the χ^2 values derived from these sets would be expected to be greater than our one actual measured value of 26.76"

From Gregory, pg. 165

If we obtain a very small p-value (e.g. a few percent?) we can interpret this as providing little support for the null hypothesis, which we may then choose to reject.

(Ultimately this choice is subjective, but χ^2 may provide objective ammunition for doing so)

If the null hypothesis were true, how probable is it that we would measure as large, or larger, a value of χ^2 ?

Recall that we refer to this important quantity as the **p-value**

$$\text{p-value} \;=\; 1 - P\left(\chi^2_{\text{obs}}\right) \;=\; 1 - \int_0^{\chi^2_{\text{obs}}} p_0\, x^{\frac{\nu}{2}-1} \exp\left(-\frac{x}{2}\right) dx \;=\; 0.02$$

What precisely does the p-value mean?

"If the galaxy flux density really *is* constant, and we repeatedly obtained sets of 15 measurements under the same conditions, then only 2% of the χ^2 values derived from these sets would be expected to be greater than our one actual measured value of 26.76"

From Gregory, pg. 165

"At this point you may be asking yourself why we should care about a probability involving results never actually obtained"

From Gregory, pg. 166

Nevertheless, p-value based frequentist hypothesis testing remains very common in the literature:

Type of problem	test	References
Line and curve goodness-of-fit	χ^2 test	NR: 15.1-15.6
Difference of means	Student's t	NR: 14.2
Ratio of variances	F test	NR: 14.2
Sample CDF	K-S test Rank sum tests	NR: 14.3, 14.6
Correlated variables?	Sample correlation coefficient	NR: 14.5, 14.6
Discrete RVs	χ^2 test / contingency table	NR: 14.4

Nevertheless, p-value based frequentist hypothesis testing remains very common in the literature:

Type of problem	test	References
Line and curve goodness-of-fit	χ^2 test	NR: 15.1-15.6
Difference of means	Student's t	NR: 14.2
Ratio of variances	F test	NR: 14.2
Sample CDF	K-S test Rank sum tests	NR: 14.3, 14.6
Correlated variables?	Sample correlation coefficient	NR: 14.5, 14.6
Discrete RVs	χ^2 test / contingency table	NR: 14.4

See also supplementary notes on my.SUPA

In the Bayesian approach, we can test our model, in the light of our data (e.g. rolling a die) and see how our knowledge of its parameters evolves, for any sample size, considering only the data that we *did* actually observe

Posterior

Likelihood

Prior

$$\boxed{p(\text{model} \,|\, \text{data}, I) \quad \propto \quad p(\text{data} \,|\, \text{model}, I) \times p(\text{model} \,|\, I)}$$

What we know now

Influence of our observations

What we knew before

In the Bayesian approach, we can test our model, in the light of our data (e.g. rolling a die) and see how our knowledge of its parameters evolves, for any sample size, considering only the data that we *did* actually observe

Posterior   Likelihood   Prior

$$\boxed{p(\text{model} \mid \text{data}, I) \quad \propto \quad p(\text{data} \mid \text{model}, I) \times p(\text{model} \mid I)}$$

What we know now   Influence of our observations   What we knew before

Astronomical example:

Probability that a galaxy is a Seyfert 1

We want to know the fraction of Seyfert galaxies which are type 1.

How large a sample do we need to reliably measure this?

We want to know the fraction of Seyfert galaxies which are type 1.

How large a sample do we need to reliably measure this?

Model as a binomial pdf: θ = global fraction of Seyfert 1s

Suppose we sample N Seyferts, and observe r Seyfert 1s

$$\boxed{p_N(r) \;\propto\; \theta^r (1-\theta)^{N-r}}$$

Likelihood = probability of obtaining observed data, given model

We want to know the fraction of Seyfert galaxies which are type 1.

How large a sample do we need to reliably measure this?

Model as a binomial pdf: θ = global fraction of Seyfert 1s

Suppose we sample N Seyferts, and observe r Seyfert 1s

$$\boxed{p_N(r) \quad \propto \quad \theta^r (1-\theta)^{N-r}}$$

Likelihood = probability of obtaining observed data, given model

Posterior Likelihood Prior

$$p(\text{model} \mid \text{data}, I) \quad \propto \quad p(\text{data} \mid \text{model}, I) \times p(\text{model} \mid I)$$

What we know now

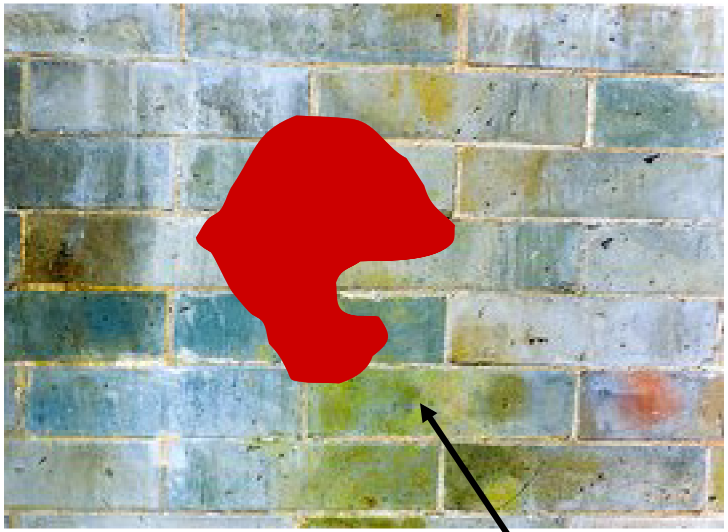
Influence of our observations

What we knew before

What do we choose as our prior?

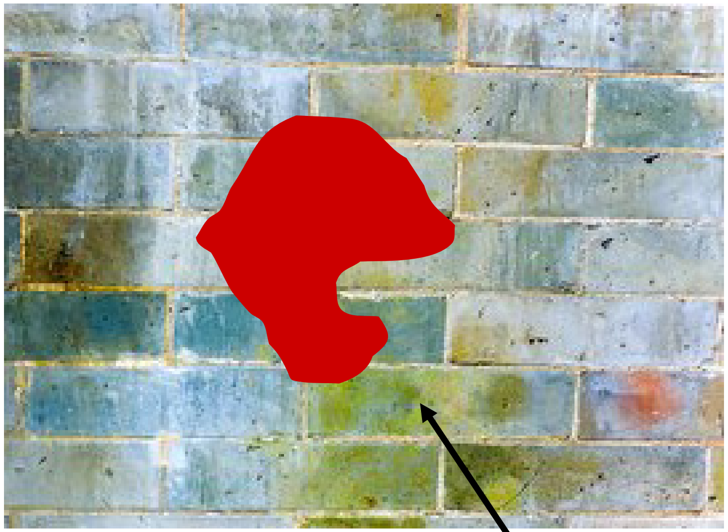
Good question!!

Source of much argument between Bayesians and frequentists



Blood on the walls

What do we choose as our prior?



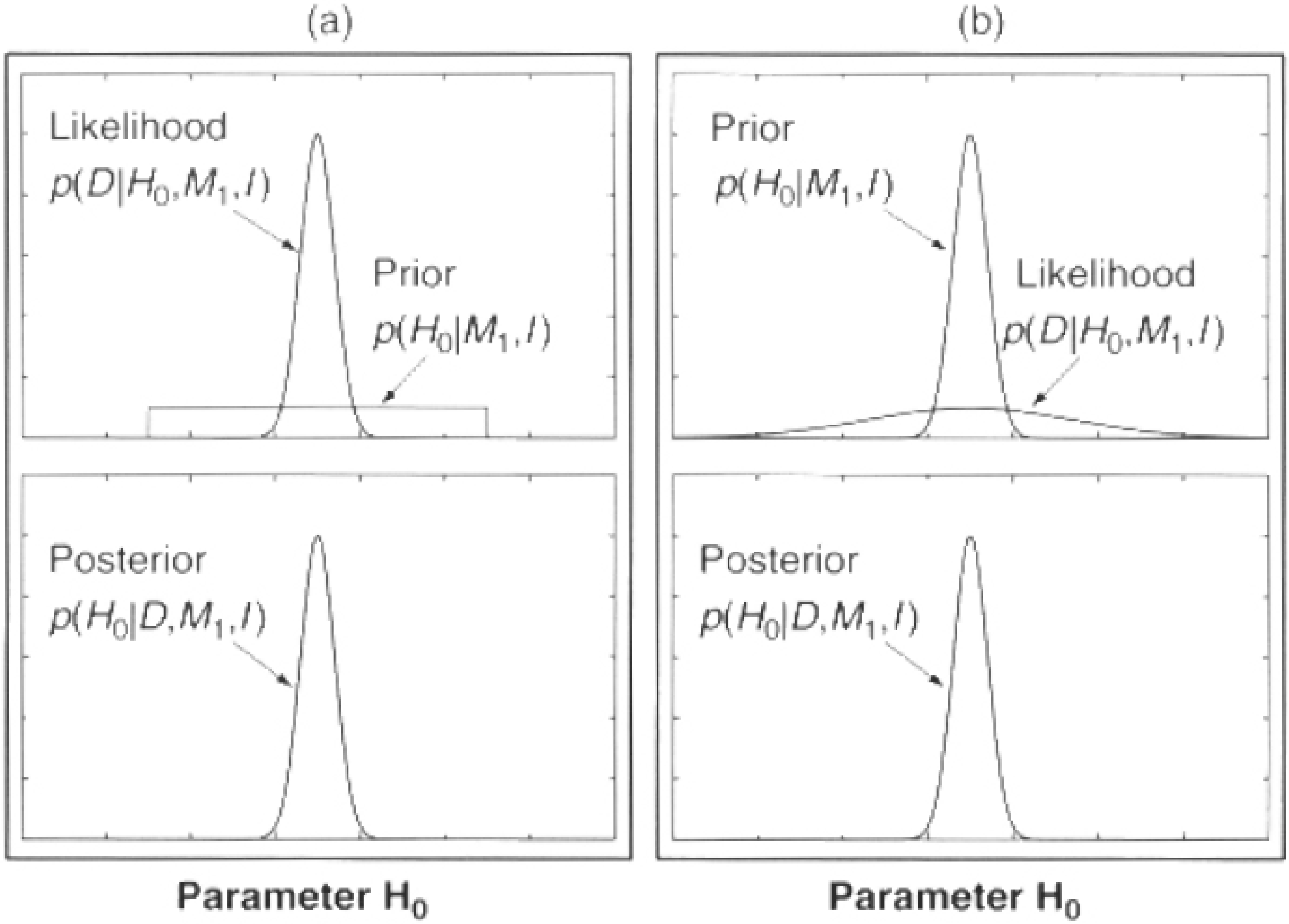
Blood on the walls

Good question!!

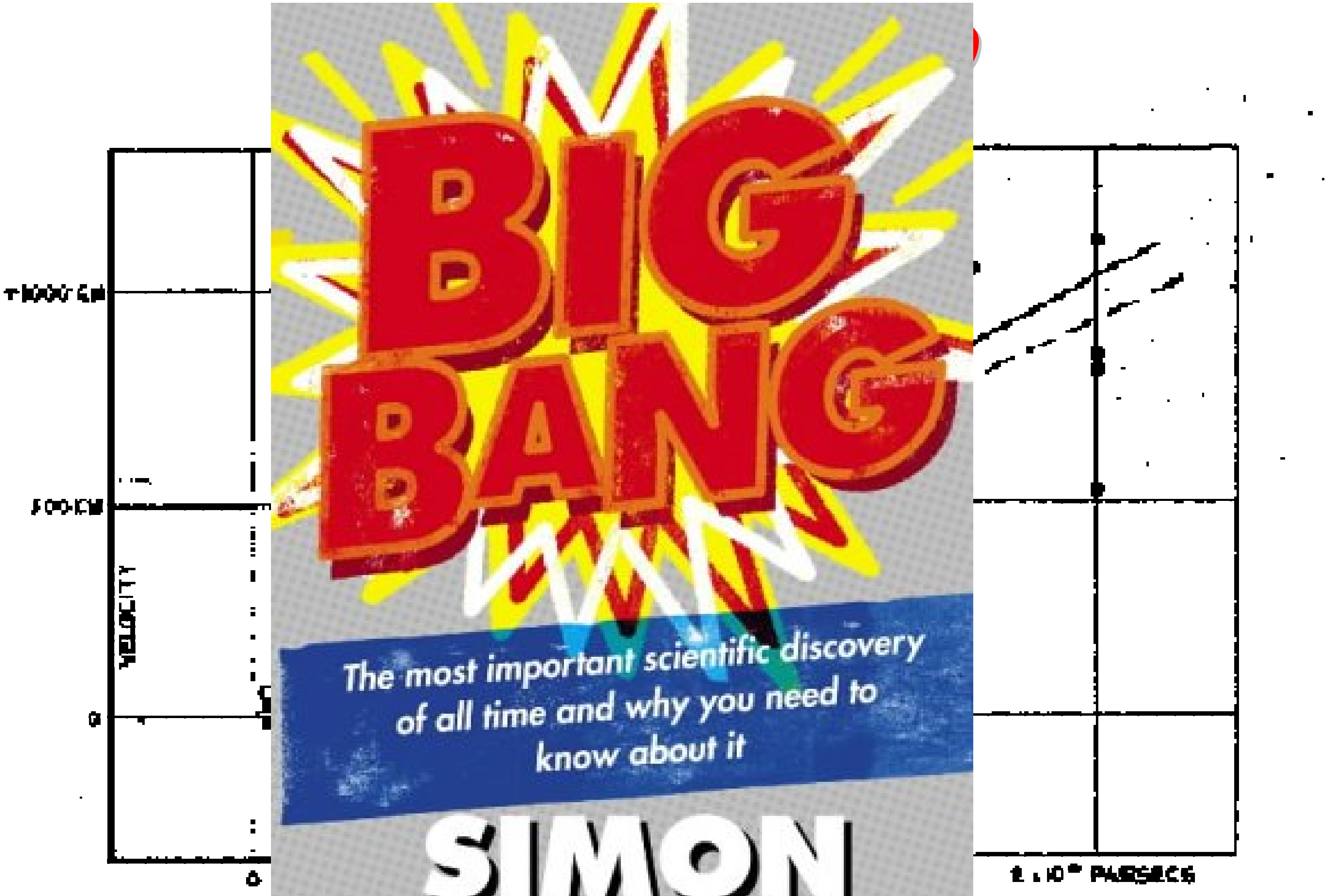
Source of much argument between Bayesians and frequentists

If our data are good enough, it shouldn't matter!!

$$\boxed{\overset{\text{Posterior}}{p(\text{model} \mid \text{data}, I)} \propto \underbrace{\overset{\text{Likelihood}}{p(\text{data} \mid \text{model}, I)}}_{\textit{Dominates}} \times \overset{\text{Prior}}{p(\text{model} \mid I)}}$$



From Gregory, pg 8.



Hubble pa ... f the Universe

'Toy' model problem: What is the relative proportion θ of Seyfert 1 galaxies in the Universe?

We can generate fake data to see how the influence of the likelihood and prior evolve.

- Choose a 'true' value of θ
- Sample a uniform random number, x, from [0,1]

  (see Numerical Recipes, and Sect 6)

3. Prob($x < \theta$) = θ

   Hence, if $x < \theta$ $\Rightarrow$ Seyfert 1

   otherwise $\Rightarrow$ Seyfert 2

4. Repeat from step 2

'Toy' model problem: What is the relative proportion θ of Seyfert 1 galaxies in the Universe?

We can generate fake data to see how the influence of the likelihood and prior evolve.

- Choose a 'true' value of θ
- Sample a uniform random number, x, from [0,1]

  (see Numerical Recipes, and Sect 6)

3. Prob($x < \theta$) = θ

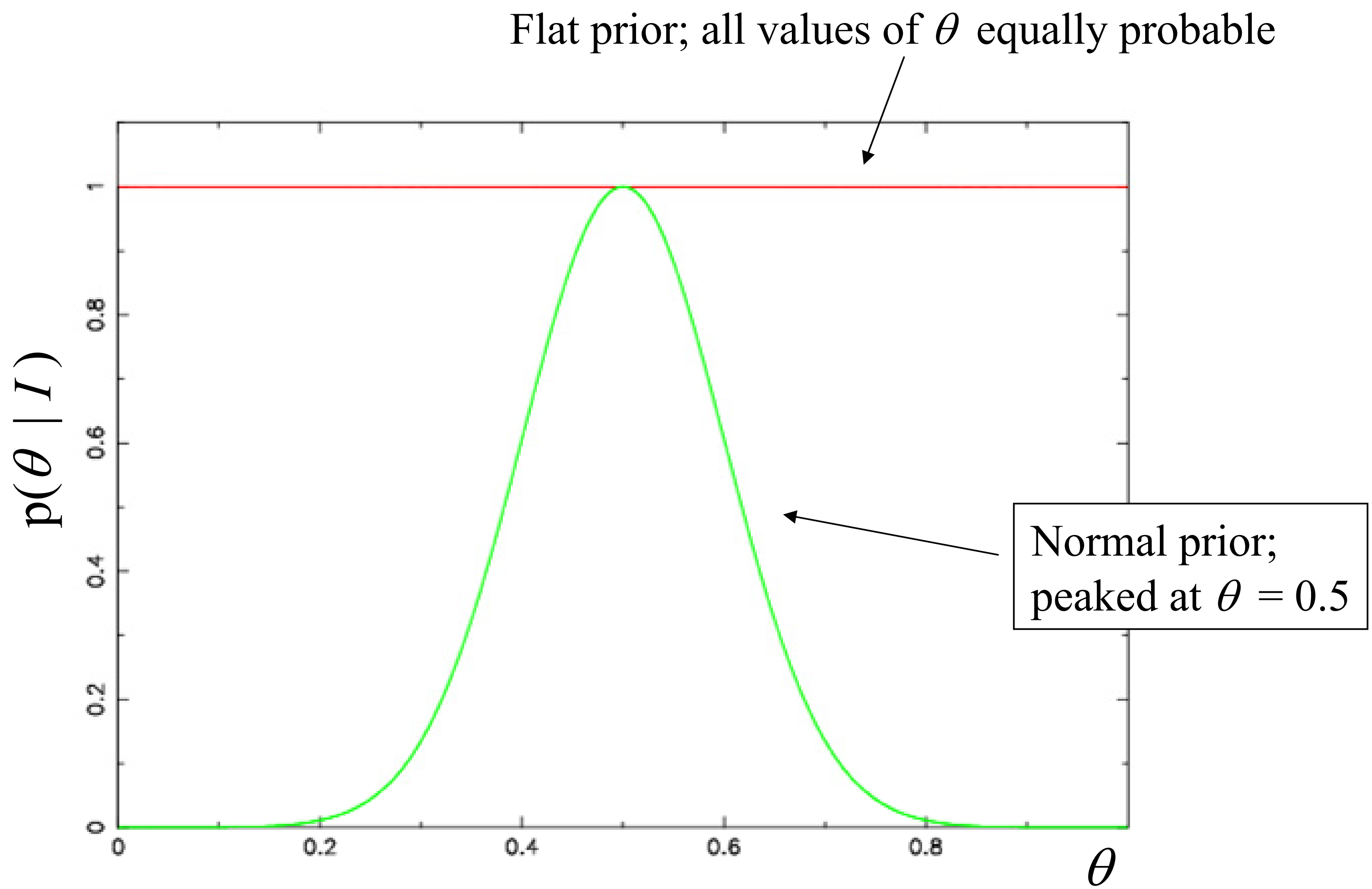
   Hence, if $x < \theta$ $\Rightarrow$ Seyfert 1

   otherwise $\Rightarrow$ Seyfert 2

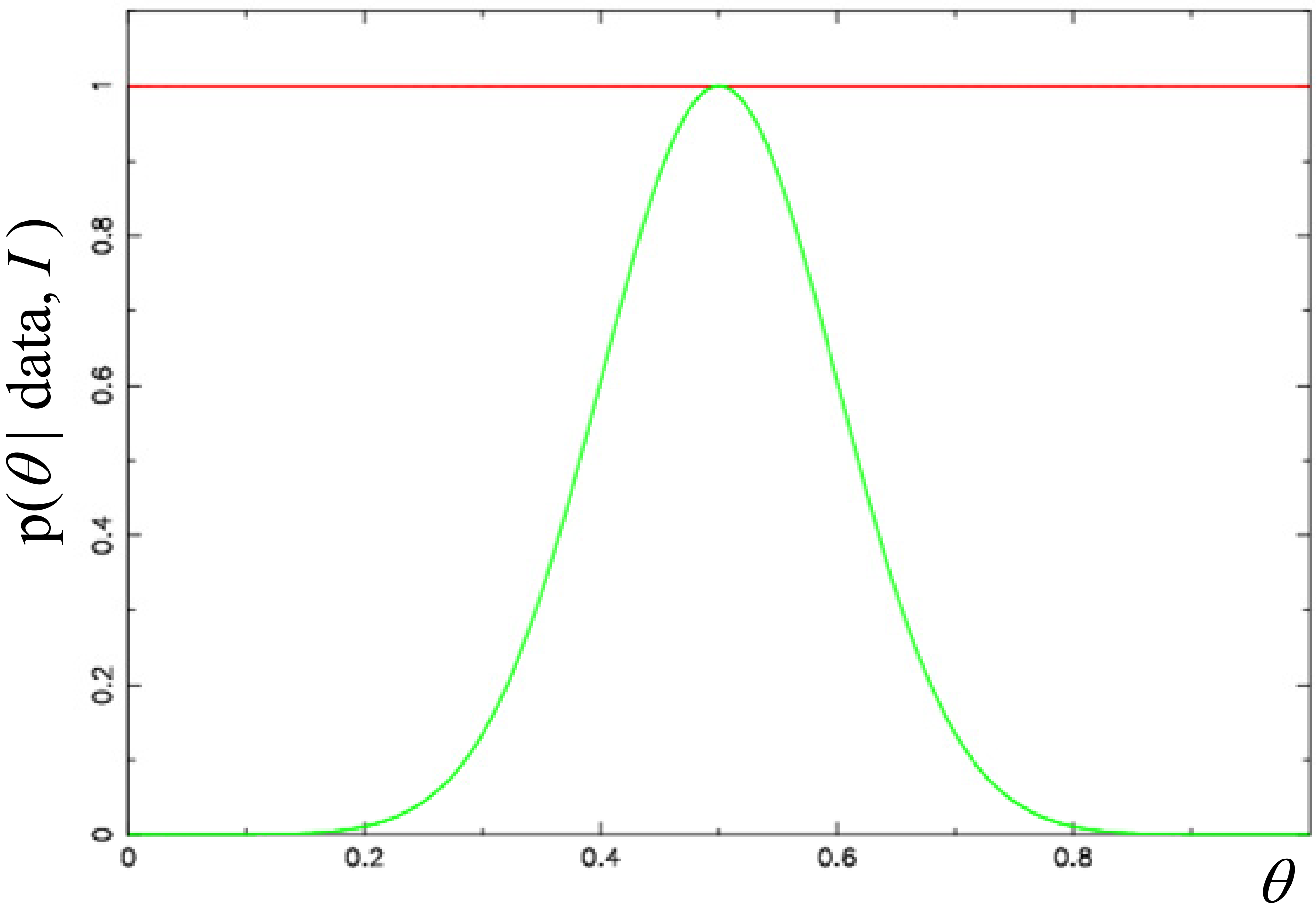
4. Repeat from step 2

Take $\theta = 0.25$

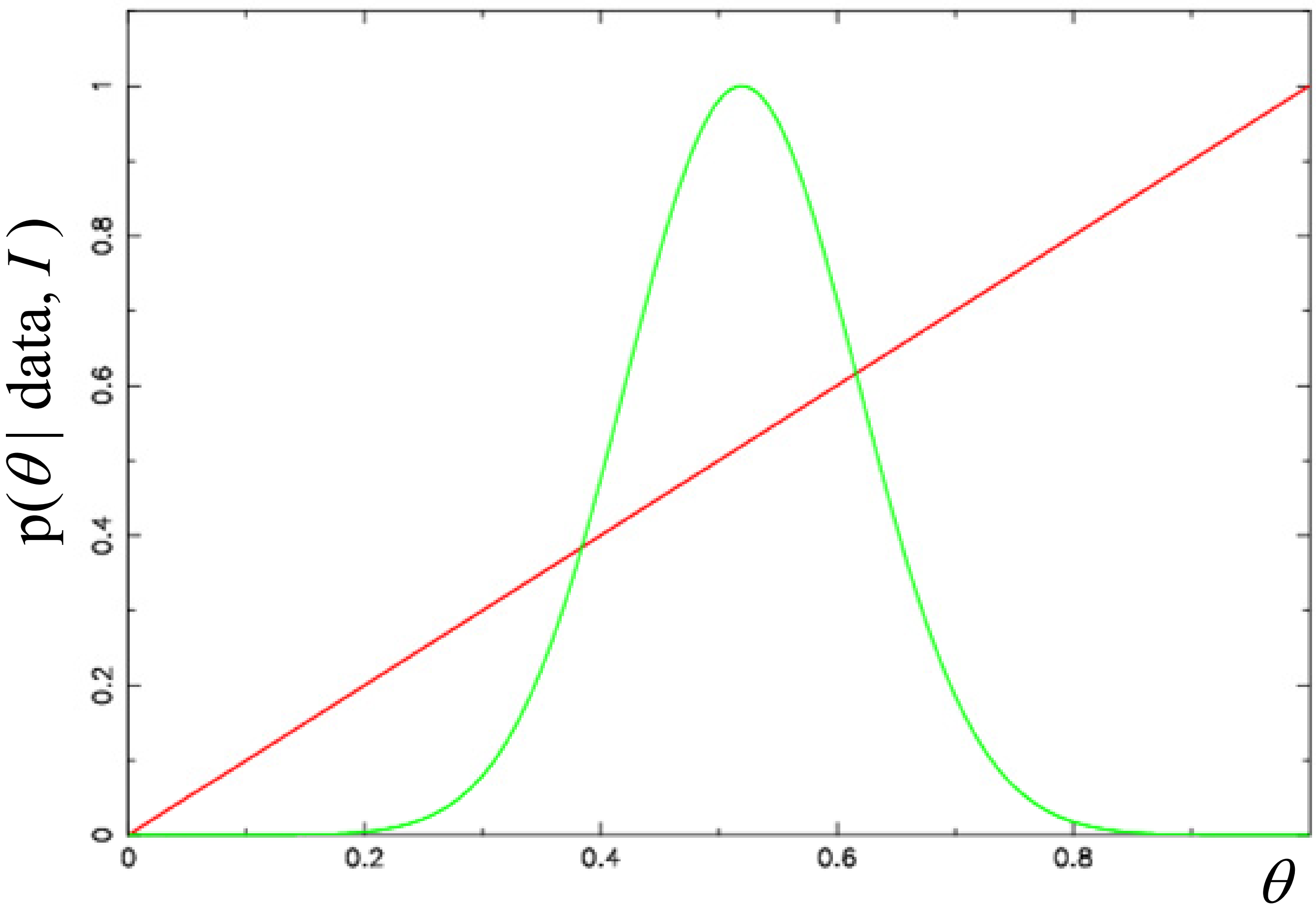
Consider two different priors



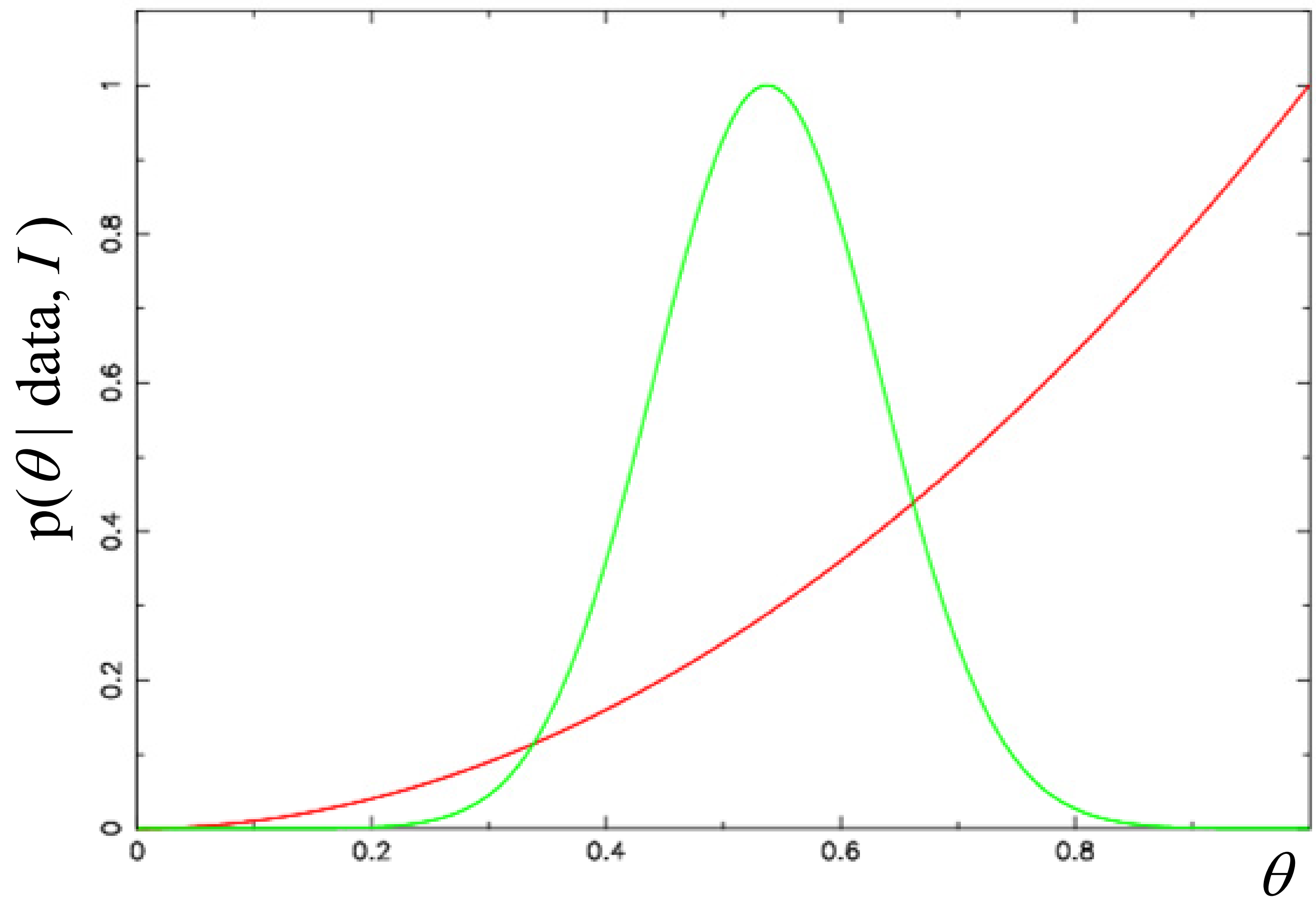
After observing 0 galaxies



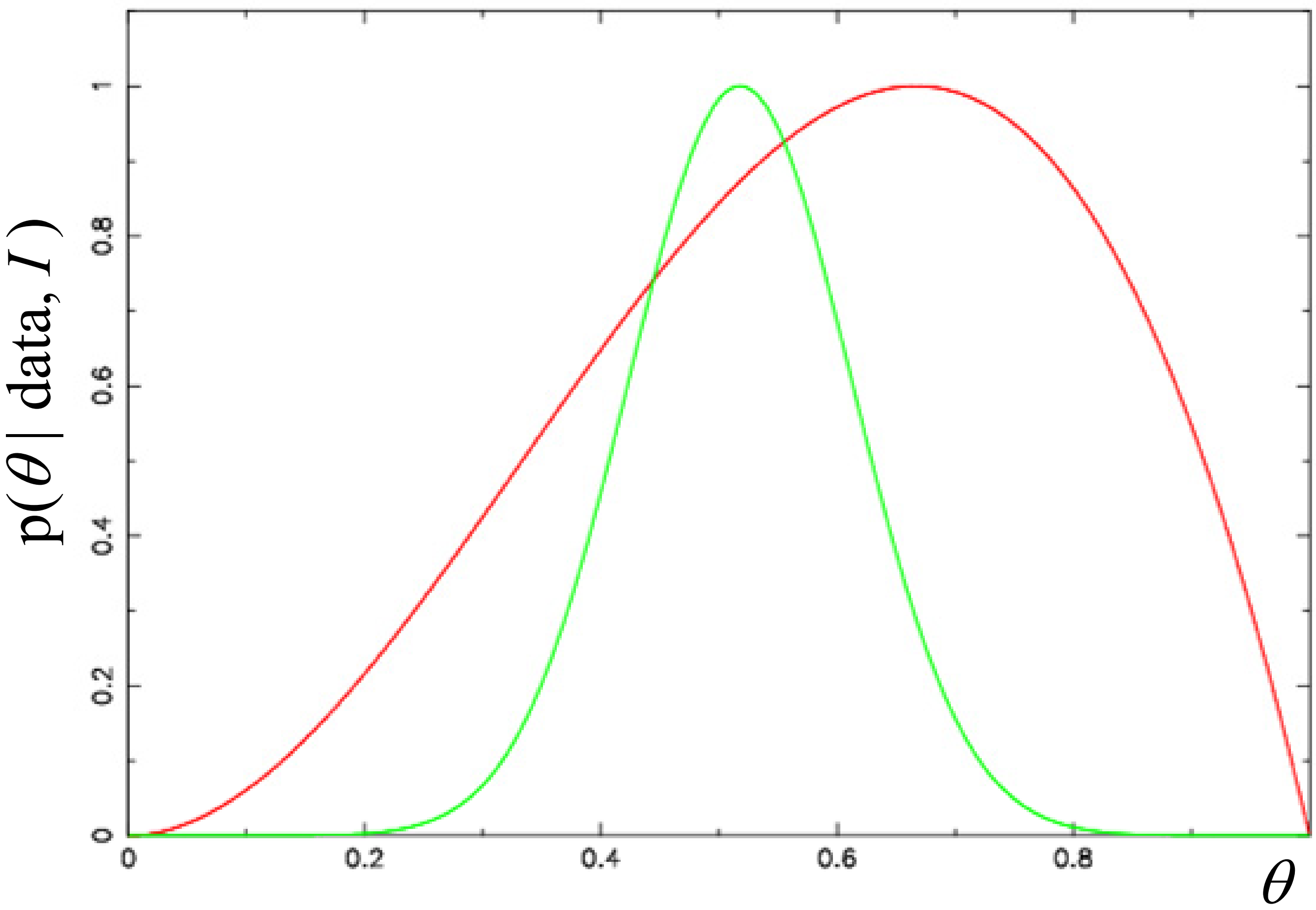
After observing **1** galaxy: Seyfert 1



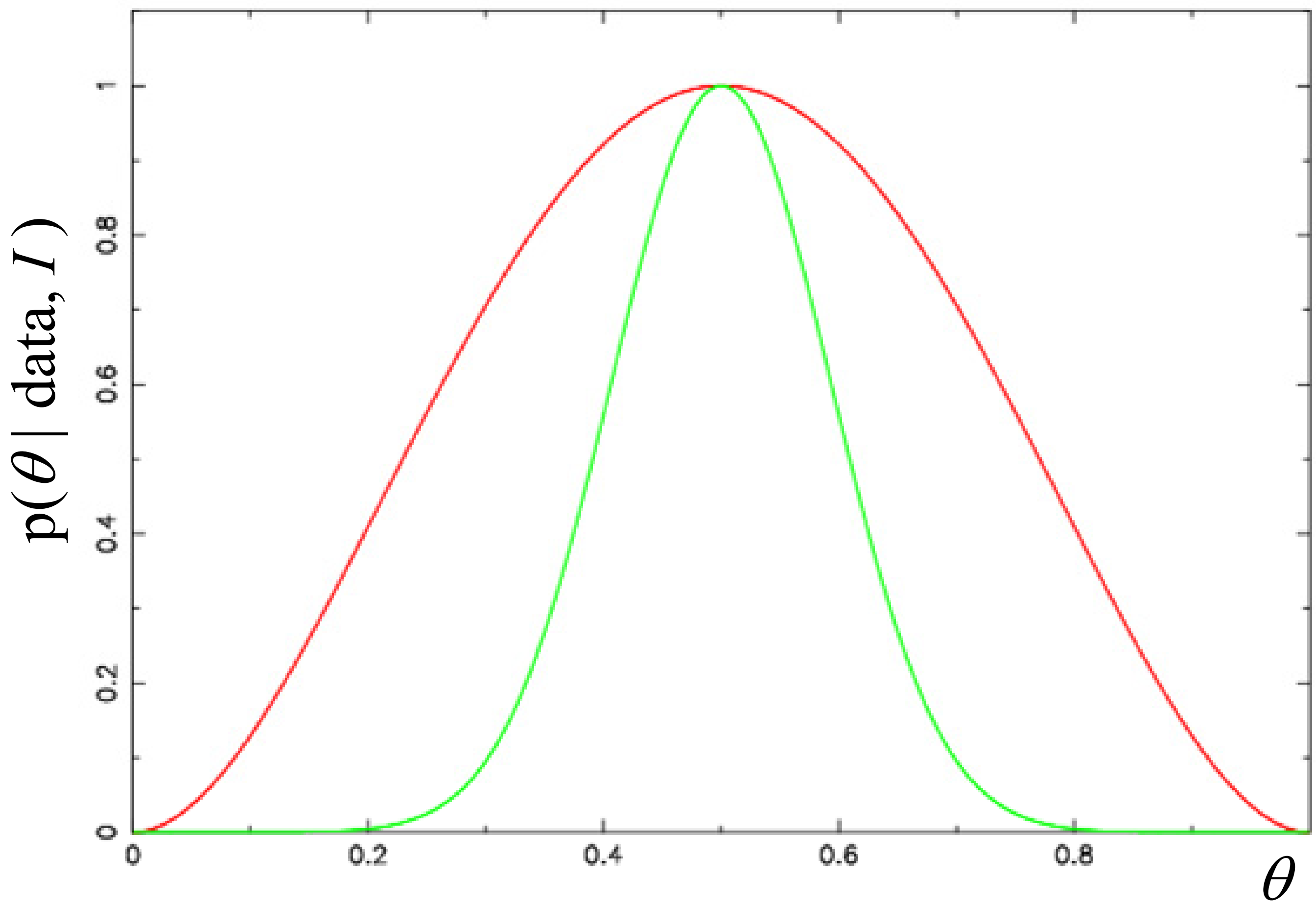
After observing 2 galaxies: S1 + S1



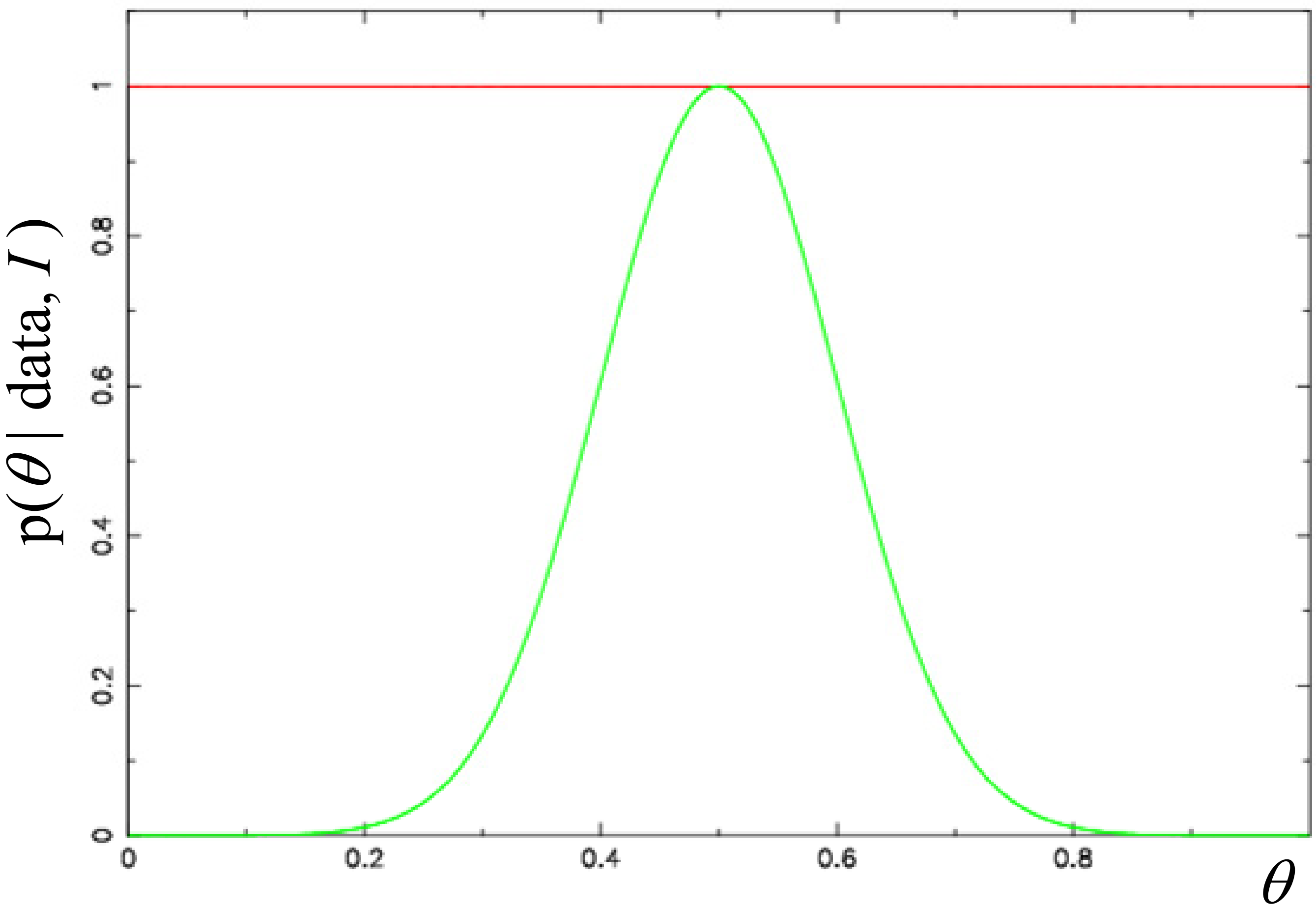
After observing **3** galaxies: S1 + S1 + S2



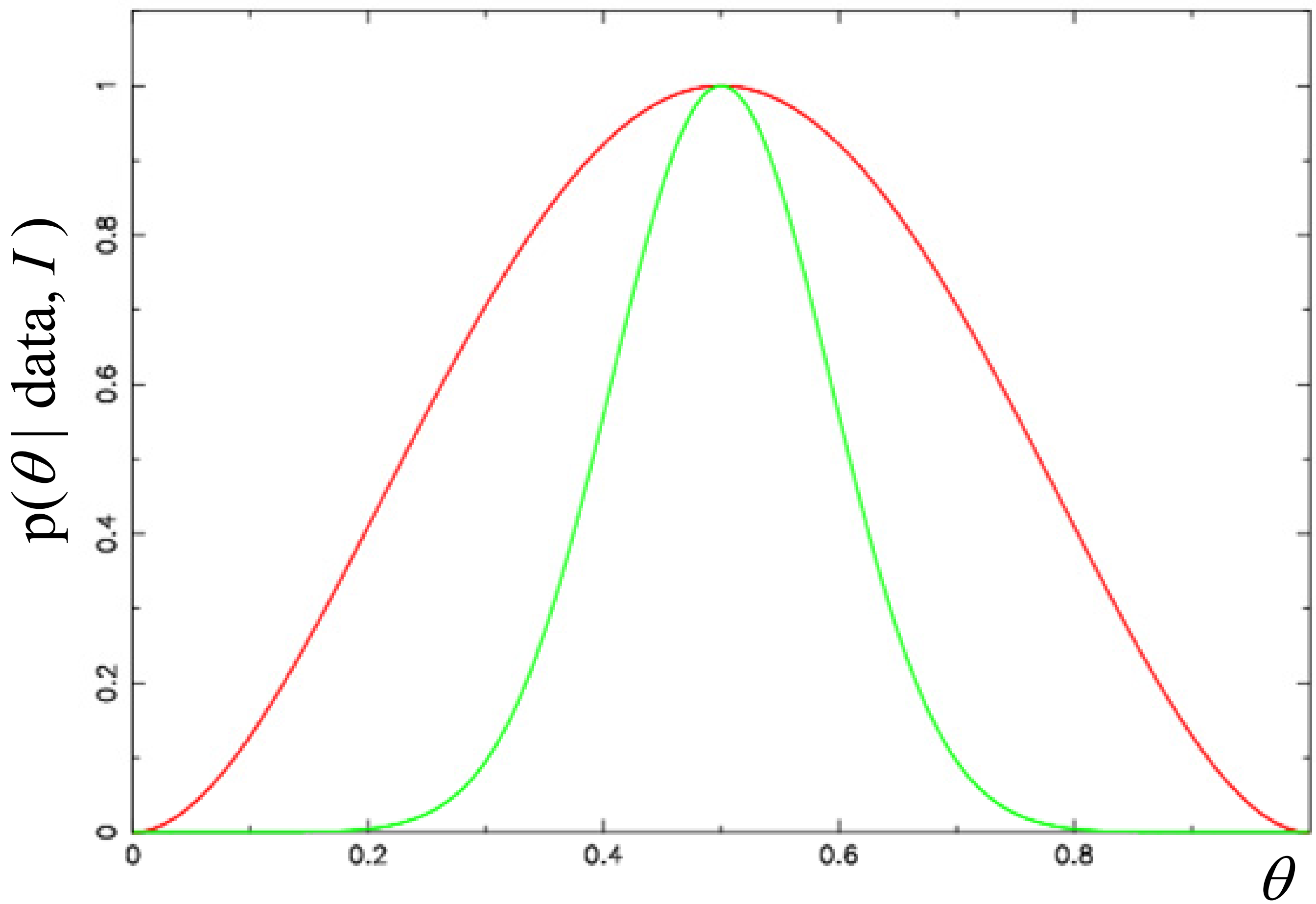
After observing **4** galaxies: S1 + S1 + S2 + S2



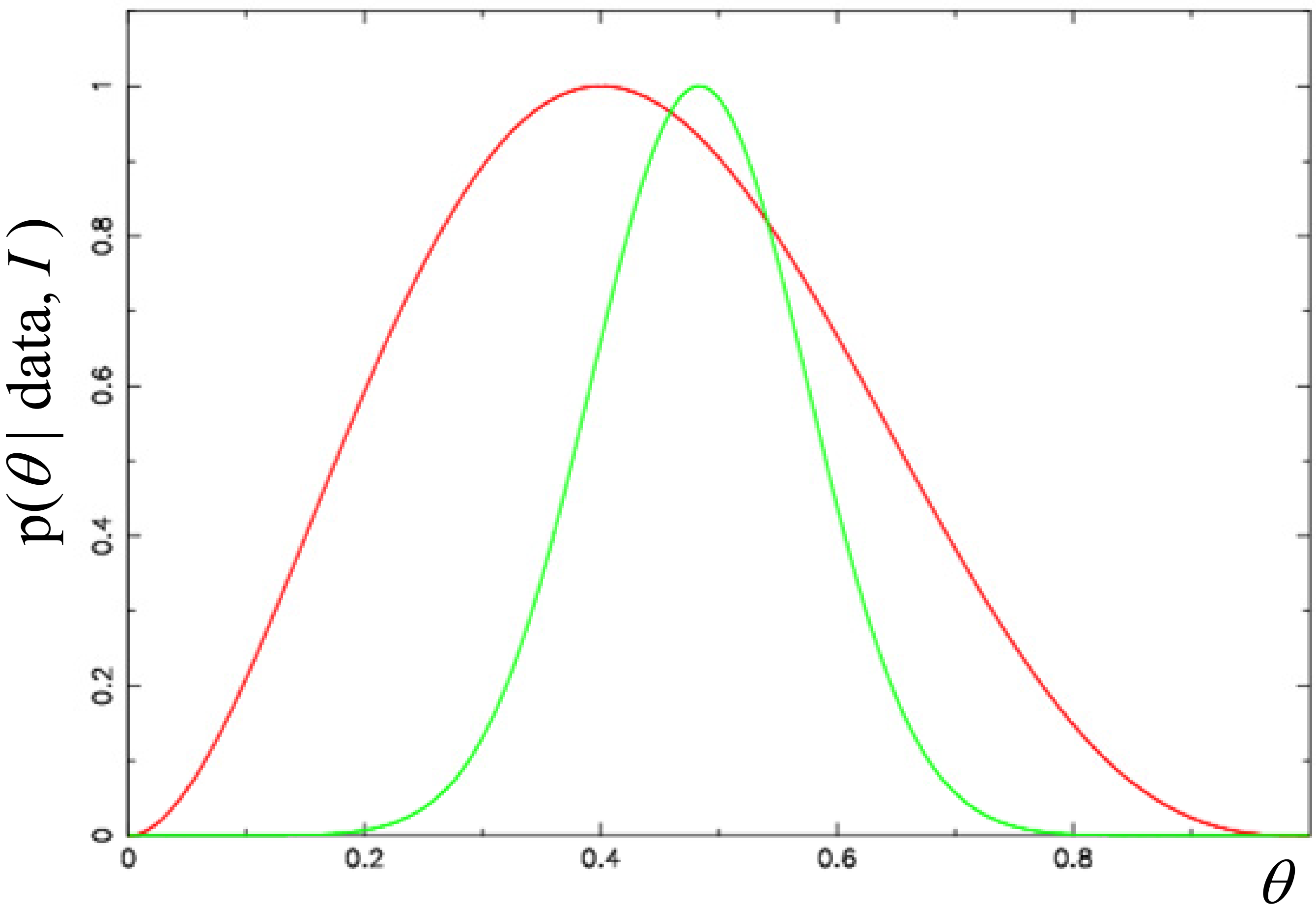
After observing 0 galaxies



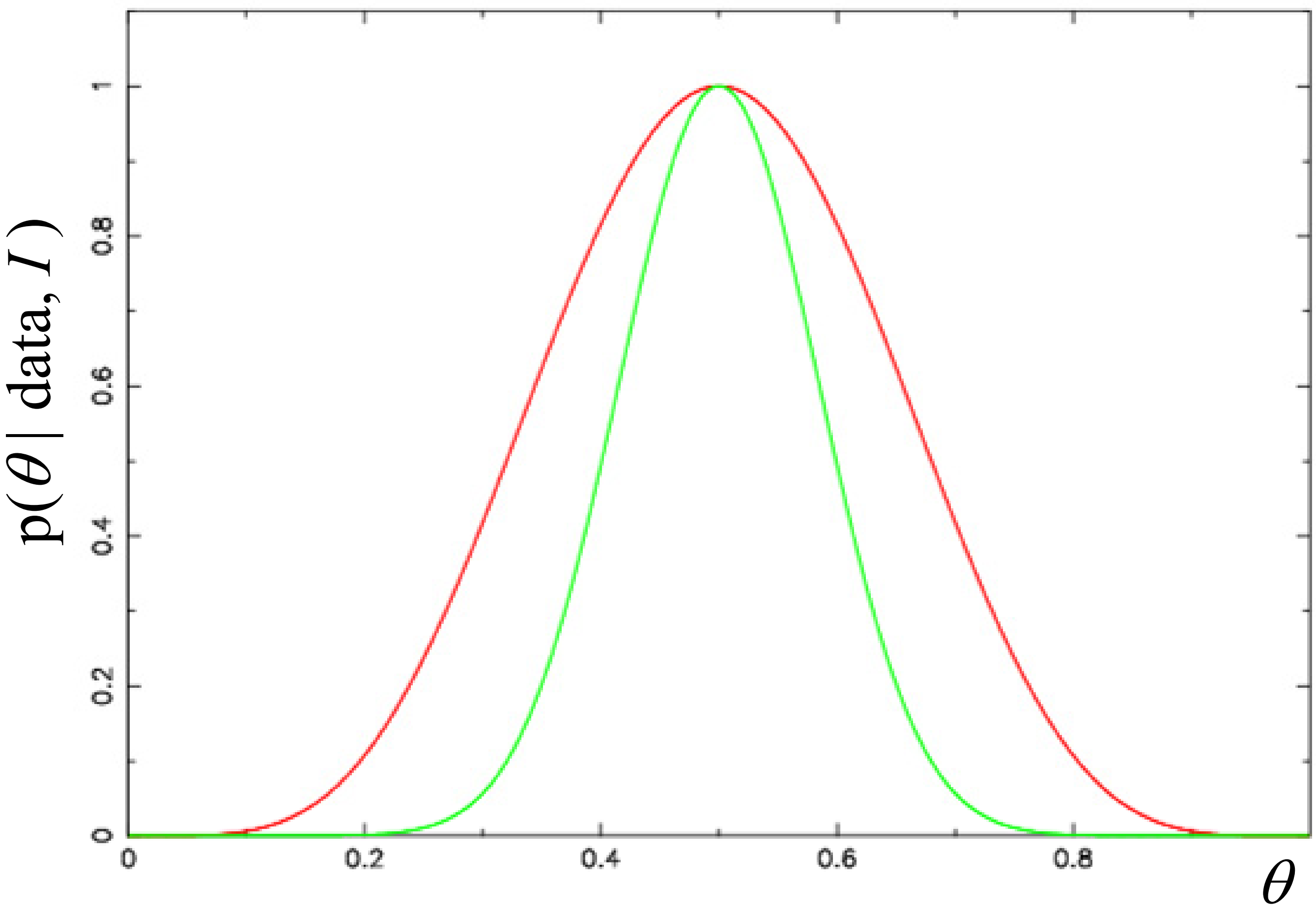
After observing **4** galaxies: S1 + S1 + S2 + S2



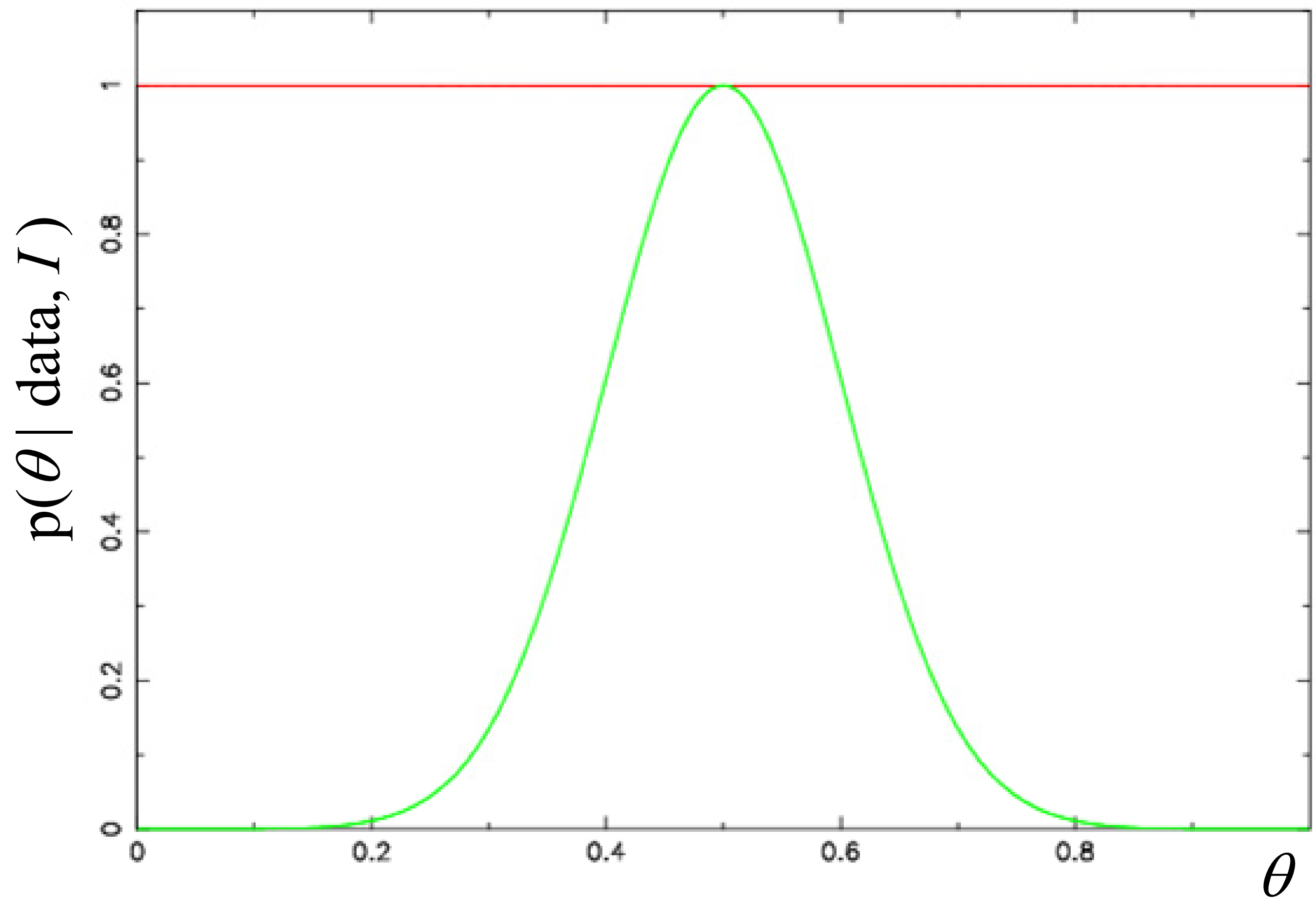
After observing **5** galaxies: S1 + S1 + S2 + S2 + S2



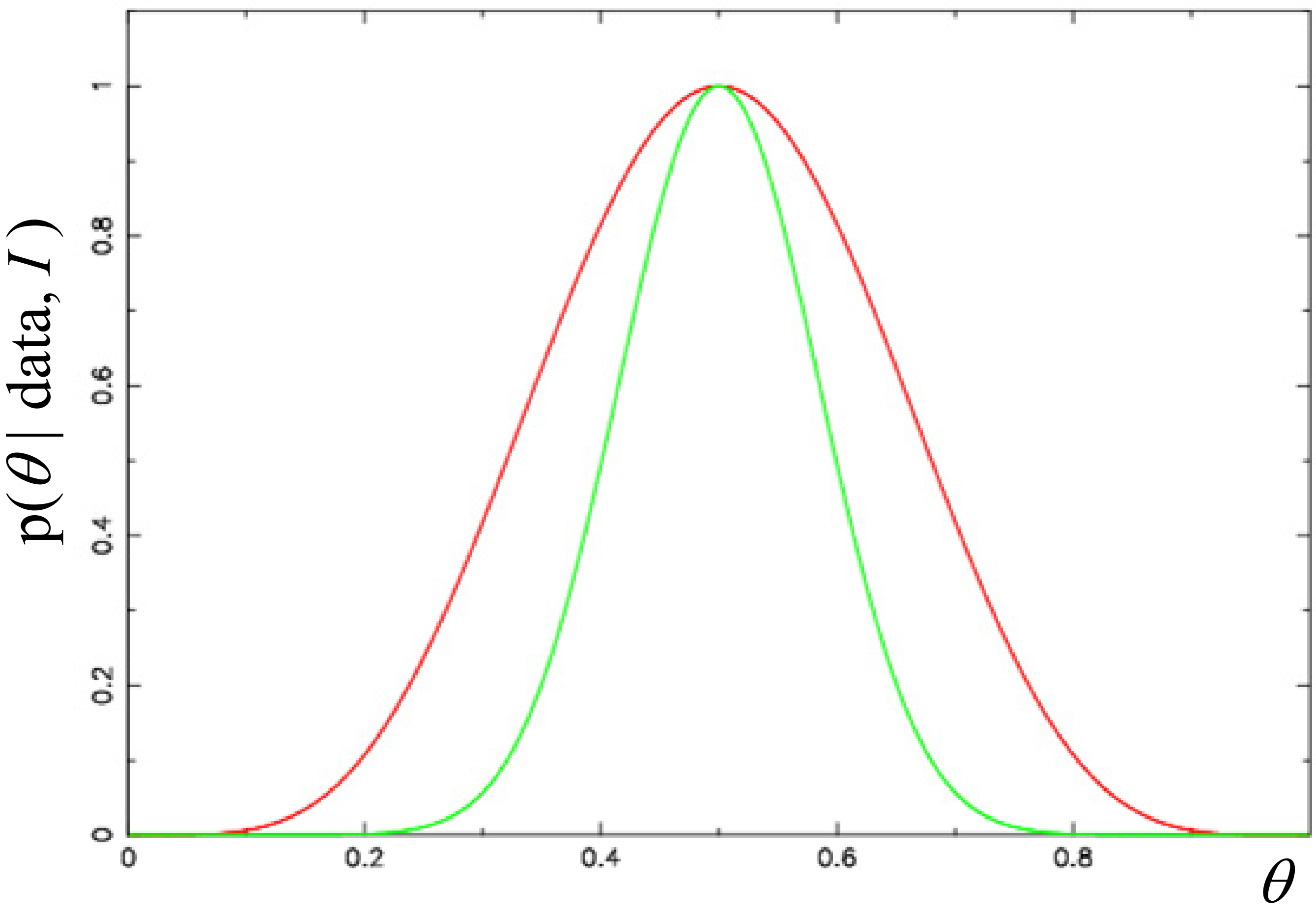
After observing **10** galaxies: 5 S1 + 5 S2



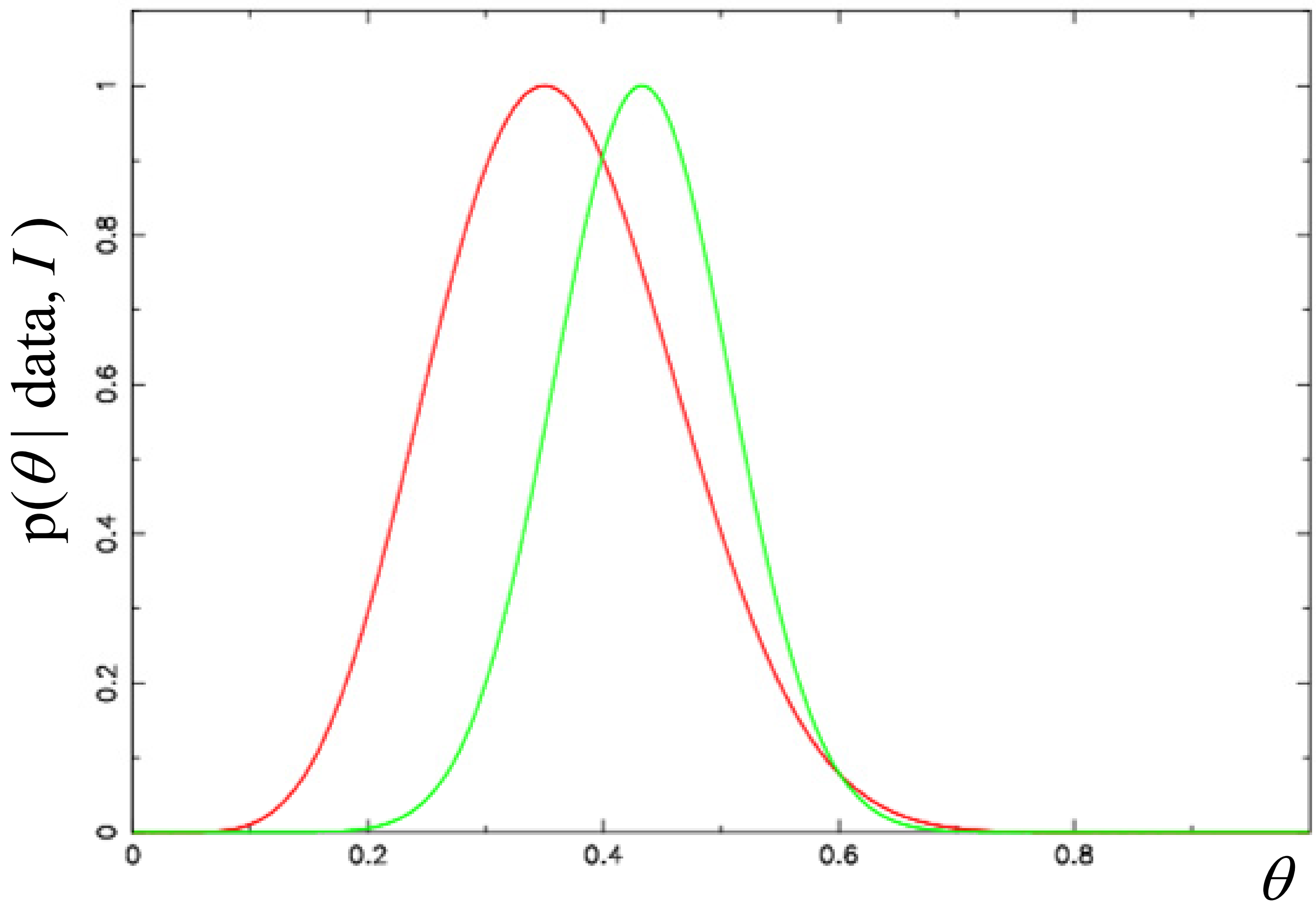
After observing 0 galaxies



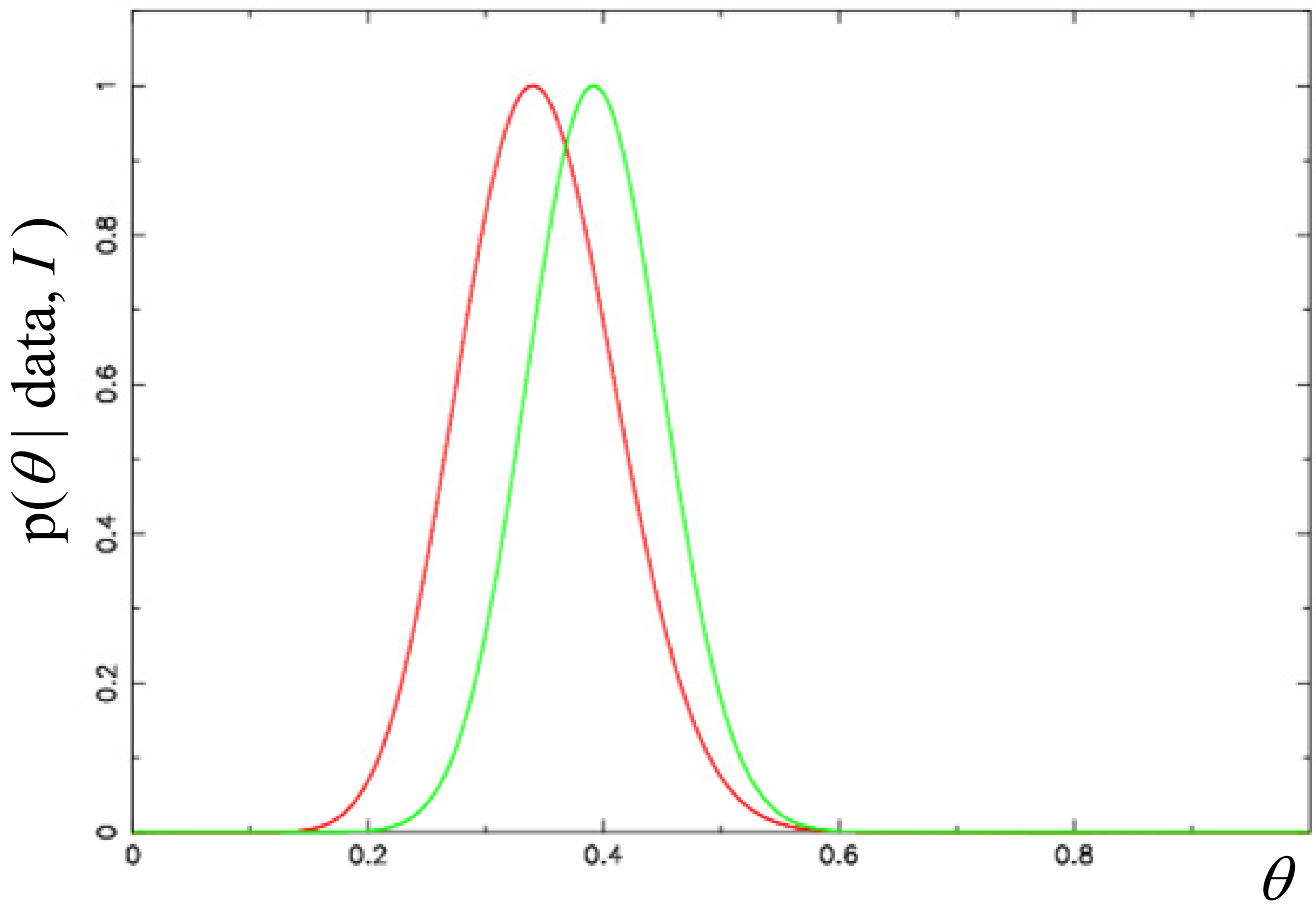
After observing **10** galaxies: 5 S1 + 5 S2



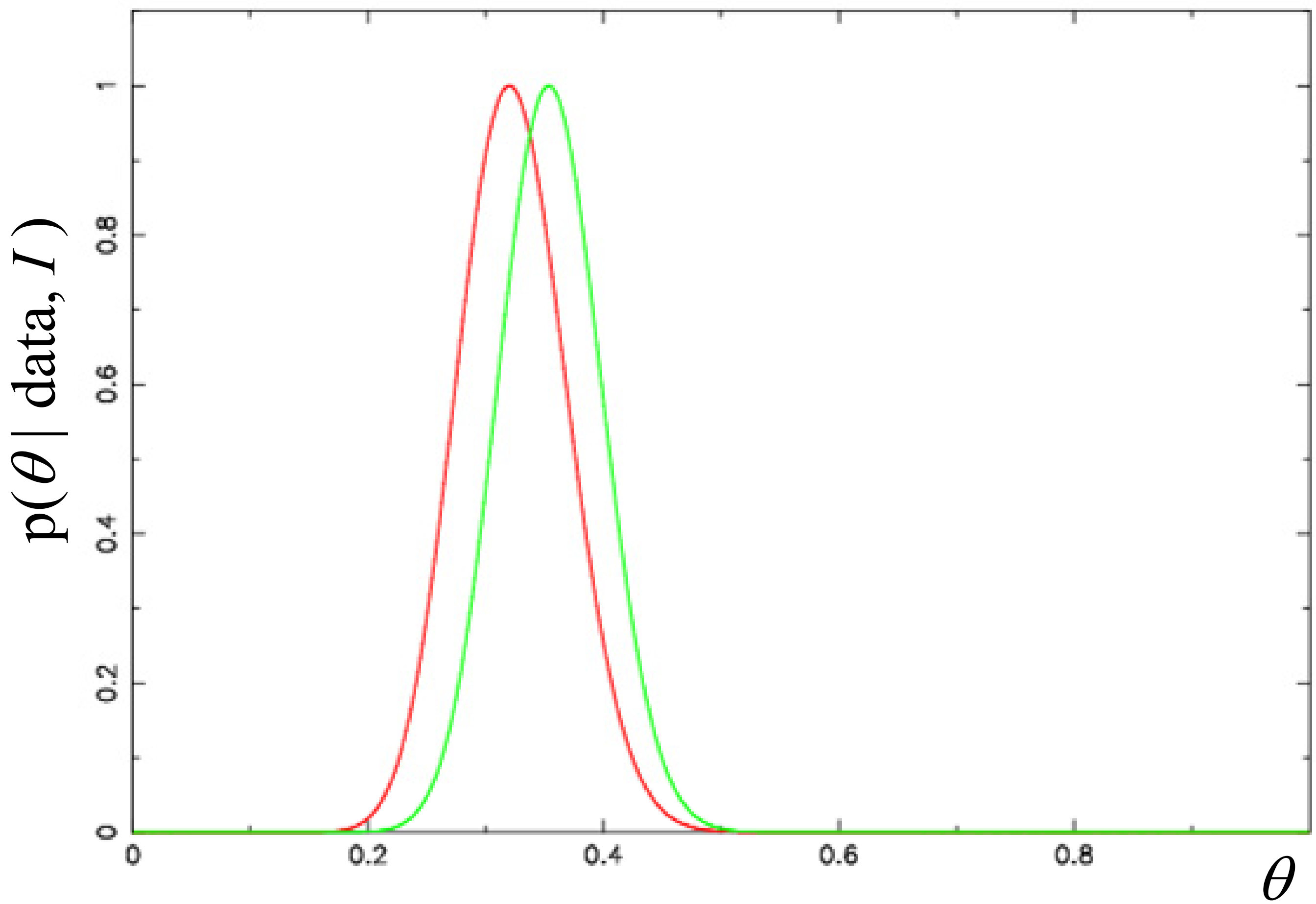
After observing 20 galaxies: 7 S1 + 13 S2



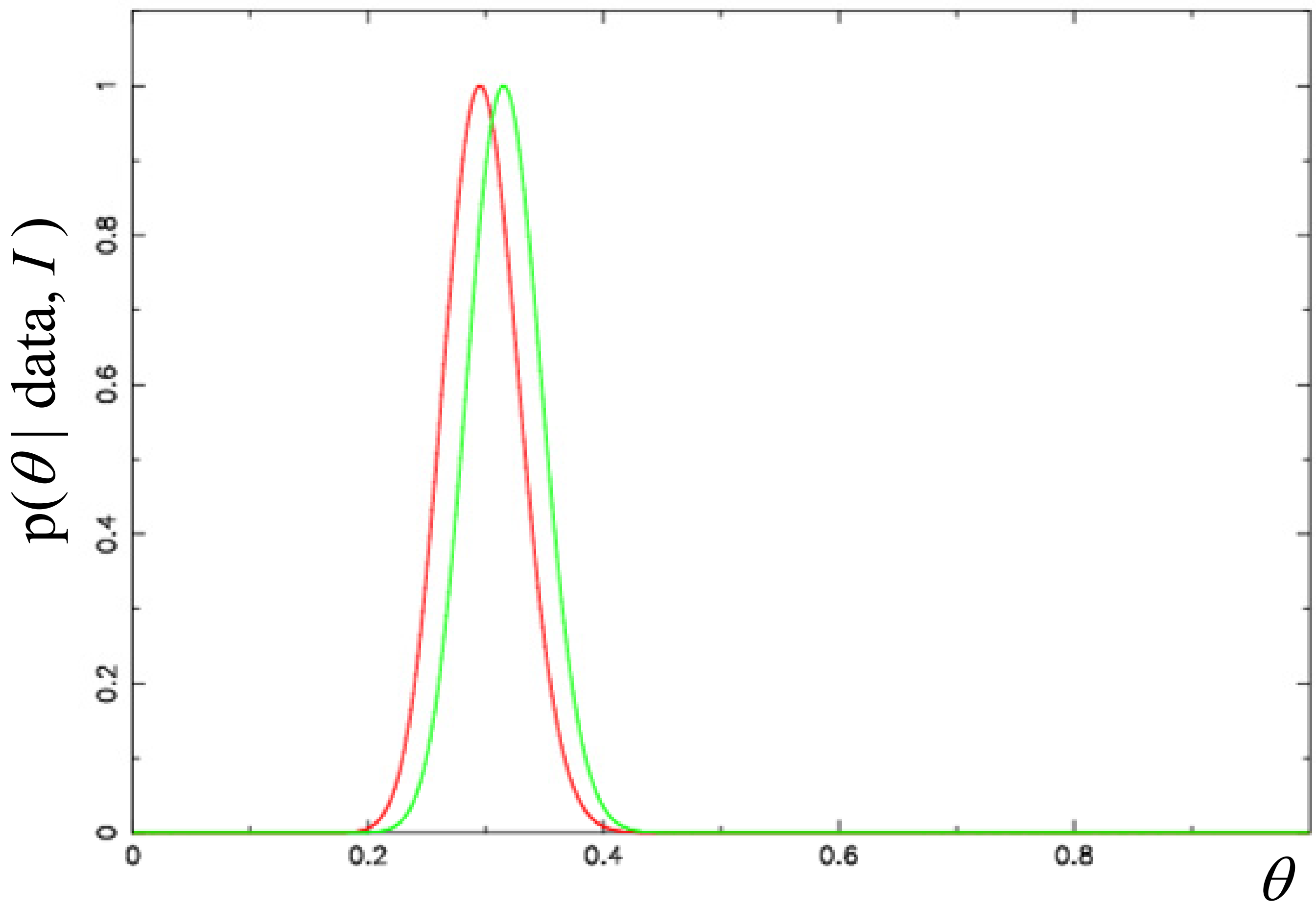
After observing **50** galaxies: 17 S1 + 33 S2



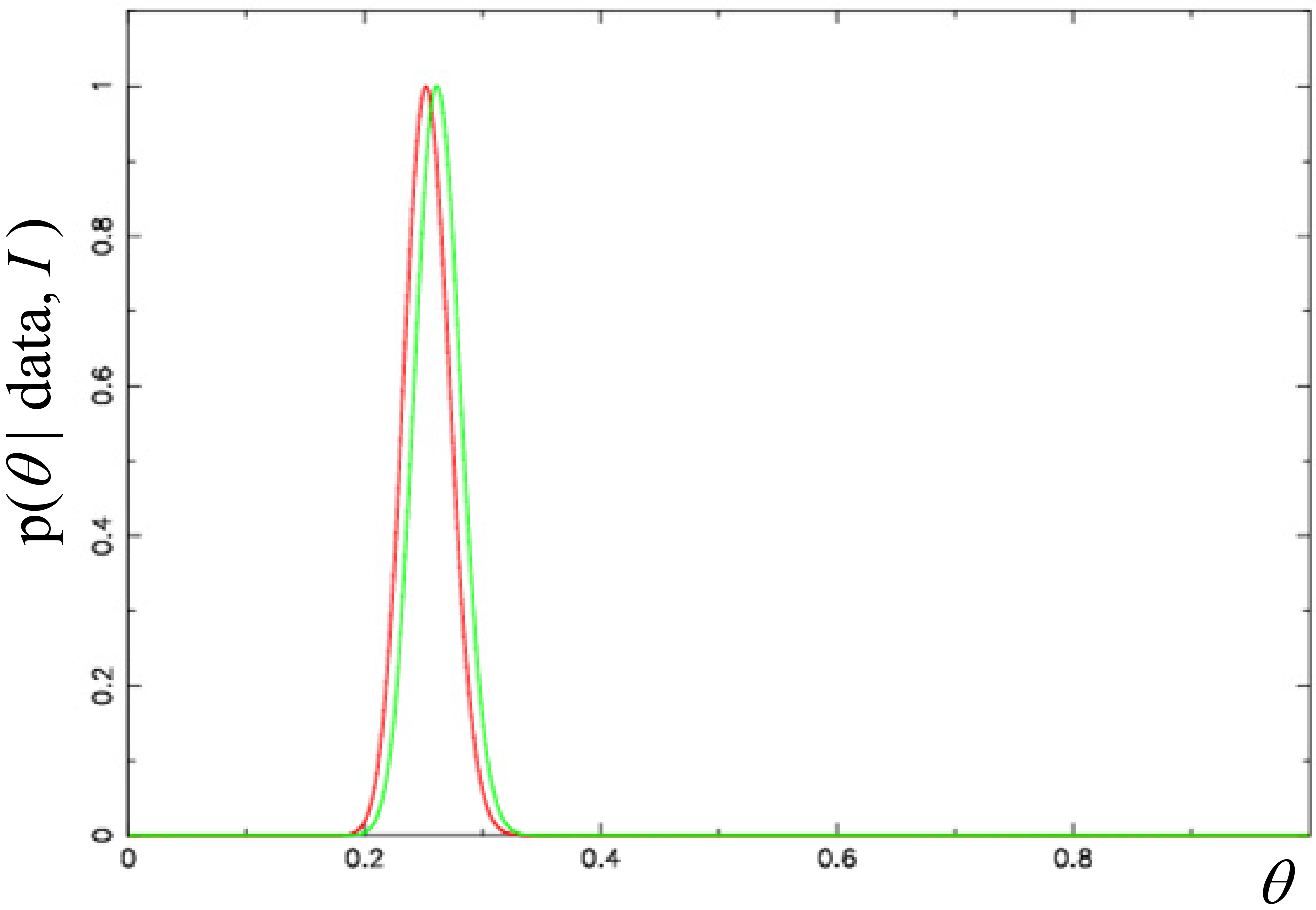
After observing **100** galaxies: 32 S1 + 68 S2



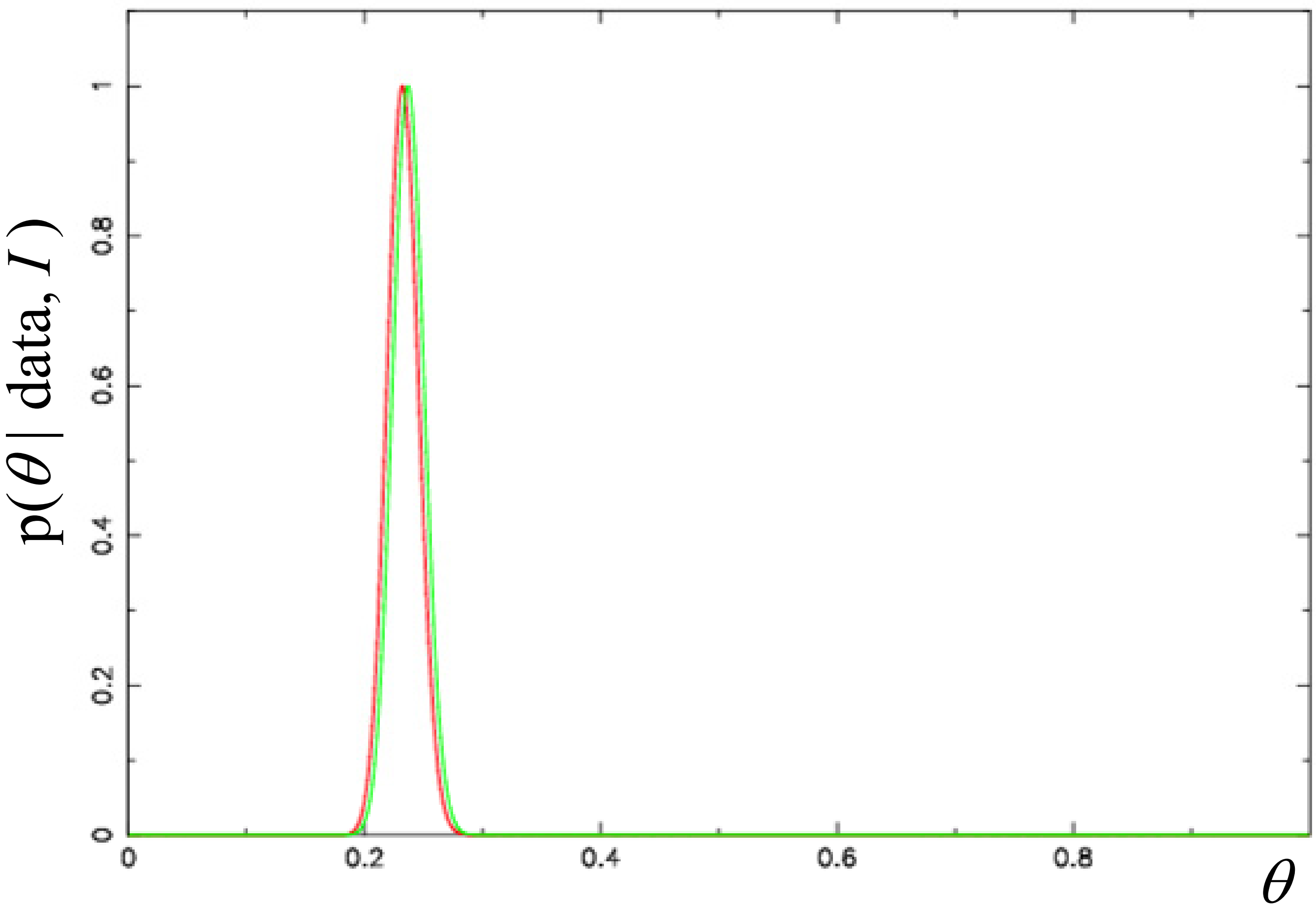
After observing **200** galaxies: 59 S1 + 141 S2



After observing 500 galaxies: 126 S1 + 374 S2



After observing 1000 galaxies: 232 S1 + 768 S2



What do we learn from all this?

- As our data improve (i.e. our sample increases), the posterior pdf narrows *and* becomes less sensitive to our choice of prior.

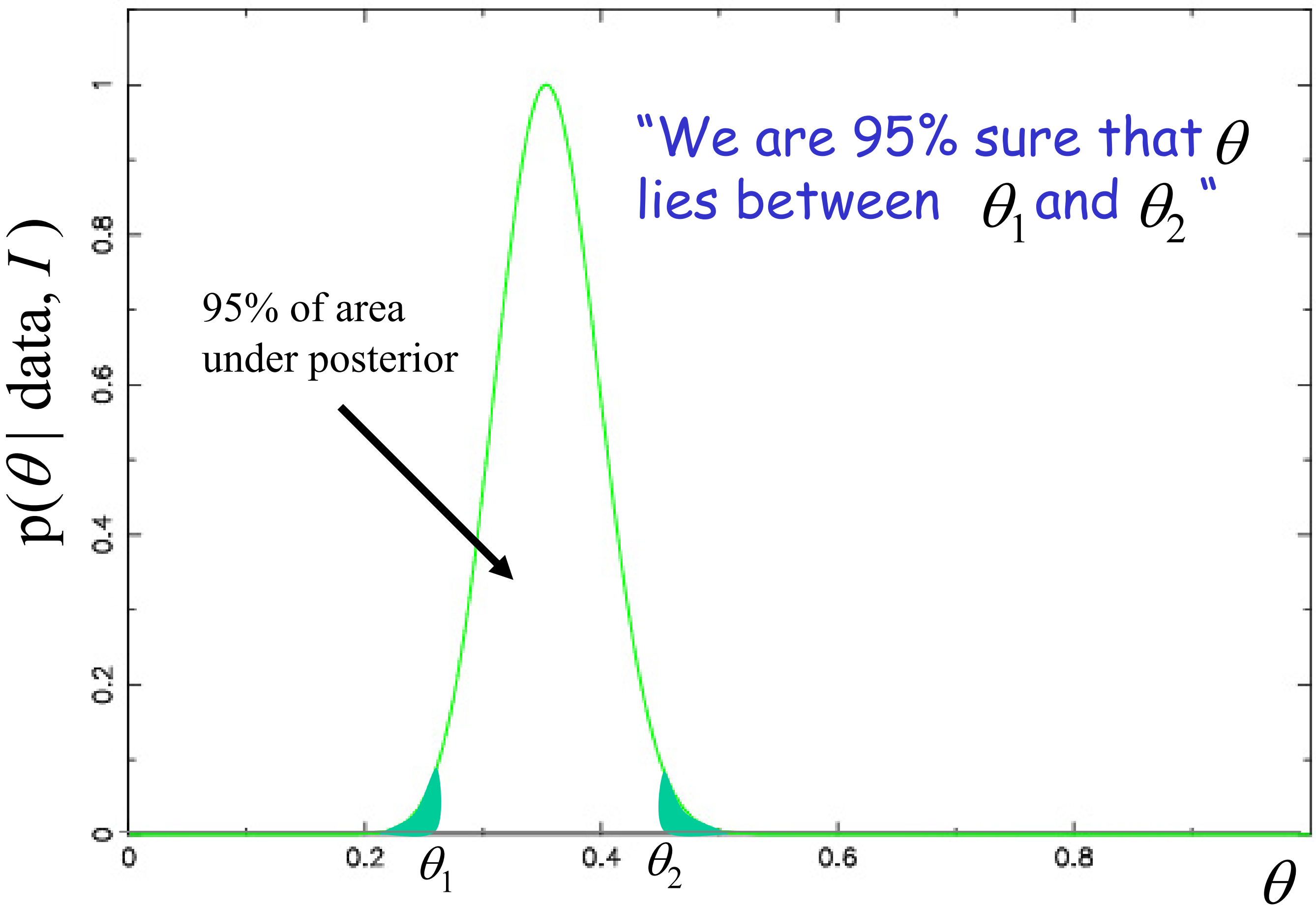
- The posterior conveys our (evolving) degree of belief in different values of θ, in the light of our data

- If we want to express our belief as a ***single number*** we can adopt e.g. the mean, median, or mode

- We can use the ***variance*** of the posterior pdf to assign an error for θ

- It is very straightforward to define Bayesian confidence intervals (more correctly termed ***credible intervals***)

Bayesian credible intervals



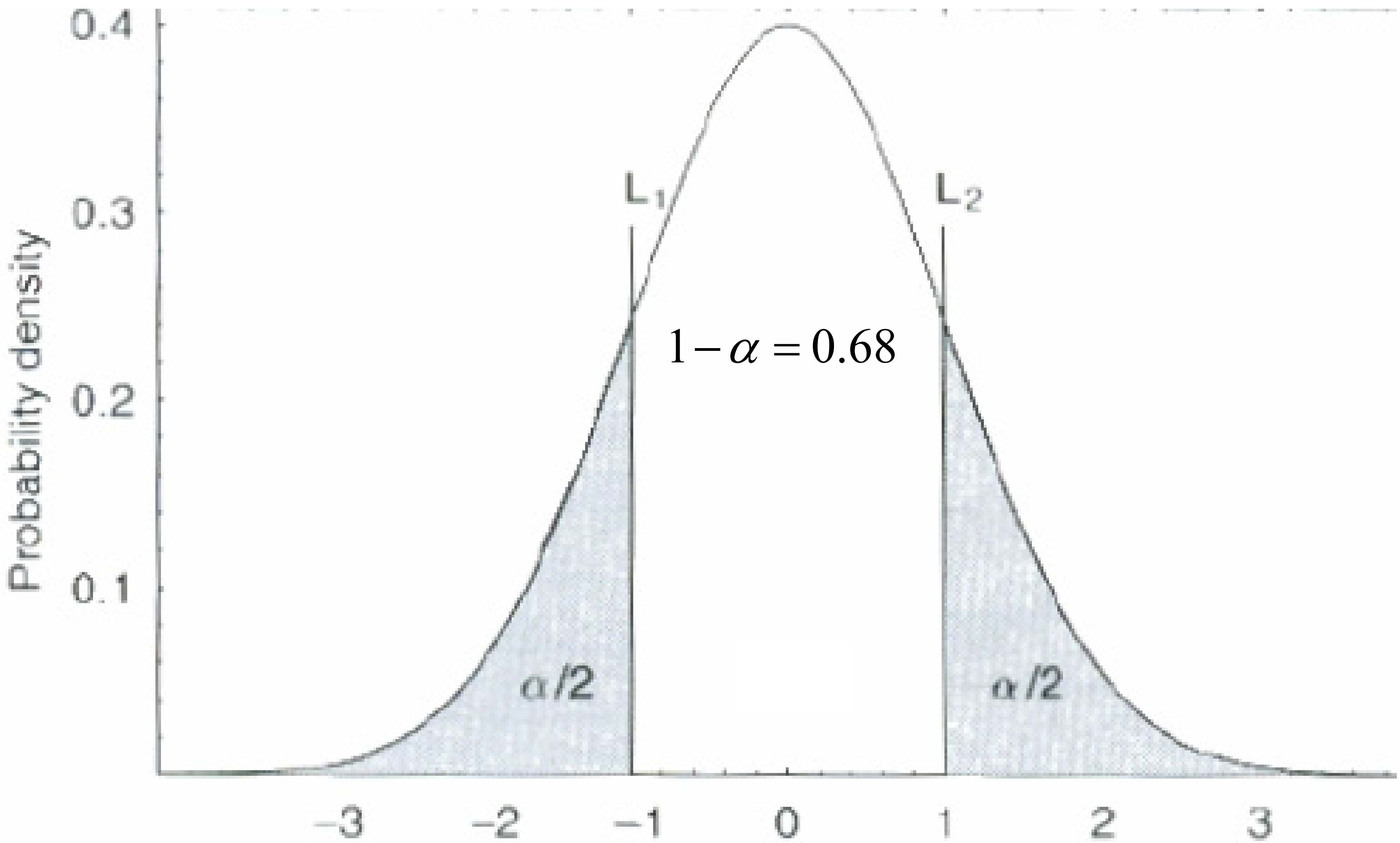
Frequentist confidence intervals

Consider an example (following Gregory pg 152)

Let $\{X_i\}$ be an iid of $n = 10$ drawn from a population $N(\mu, \sigma^2)$ with unknown μ but known $\sigma = 1$.

Let $\bar{X}$ be the sample mean RV, which has SD $\sigma_m = \sigma / \sqrt{10} \sim 0.32$

Thus $$\boxed{\text{Prob}\left(\mu - 0.32 < \bar{X} < \mu + 0.32\right) = 0.68}$$



$$z = \frac{\overline{X} - \mu}{\sigma_m}$$

Frequentist confidence intervals

Consider an example (following Gregory pg 152)

Let $\{X_i\}$ be an iid of $n=10$ drawn from a population $N(\mu,\sigma^2)$ with unknown μ but known $\sigma=1$.

Let $\overline{X}$ be the sample mean RV, which has SD $\sigma_m = \sigma/\sqrt{10} \sim 0.32$

Thus $$\boxed{\text{Prob}\left(\mu - 0.32 < \overline{X} < \mu + 0.32\right) = 0.68}$$

We can re-arrange this to write

$$\boxed{\text{Prob}\left(\overline{X} - 0.32 < \mu < \overline{X} + 0.32\right) = 0.68}$$

Suppose that from our observed sample we measure $\bar{x} = 5.40$

Can we simply write $\text{Prob}\left(5.08 < \mu < 5.72\right) = 0.68$?

Suppose that from our observed sample we measure $\bar{x} = 5.40$

Can we simply write $\text{Prob}\left(5.08 < \mu < 5.72\right) = 0.68$?

No!

Question 8: We can't write $\text{Prob}\left(5.08 < \mu < 5.72\right) = 0.68$ because μ is a fixed (but unknown) parameter. Hence, which of the following statements is true?

A $\quad \text{Prob}\left(5.08 < \mu < 5.72\right) \neq 0$

B $\quad \text{Prob}\left(5.08 < \mu < 5.72\right) \neq 1$

C $\quad 0 < \text{Prob}\left(5.08 < \mu < 5.72\right) < 1$

D $\quad \text{Prob}\left(5.08 < \mu < 5.72\right) = 0$ or 1

Question 8: We can't write $\mathrm{Prob}\left(5.08 < \mu < 5.72\right) = 0.68$ because μ is a fixed (but unknown) parameter. Hence, which of the following statements is true?

A $\mathrm{Prob}\left(5.08 < \mu < 5.72\right) \neq 0$

B $\mathrm{Prob}\left(5.08 < \mu < 5.72\right) \neq 1$

C $0 < \mathrm{Prob}\left(5.08 < \mu < 5.72\right) < 1$

D $\mathrm{Prob}\left(5.08 < \mu < 5.72\right) = 0$ or 1

Suppose that from our observed sample we measure $\bar{x} = 5.40$

Can we simply write $\text{Prob}(5.08 < \mu < 5.72) = 0.68$?

No!

In the frequentist approach, the true mean μ is a fixed (although unknown) parameter – it either belongs to the interval (5.08,5.72) or it doesn't! Thus

$$\boxed{\text{Prob}(5.08 < \mu < 5.72) = 0 \text{ or } 1}$$

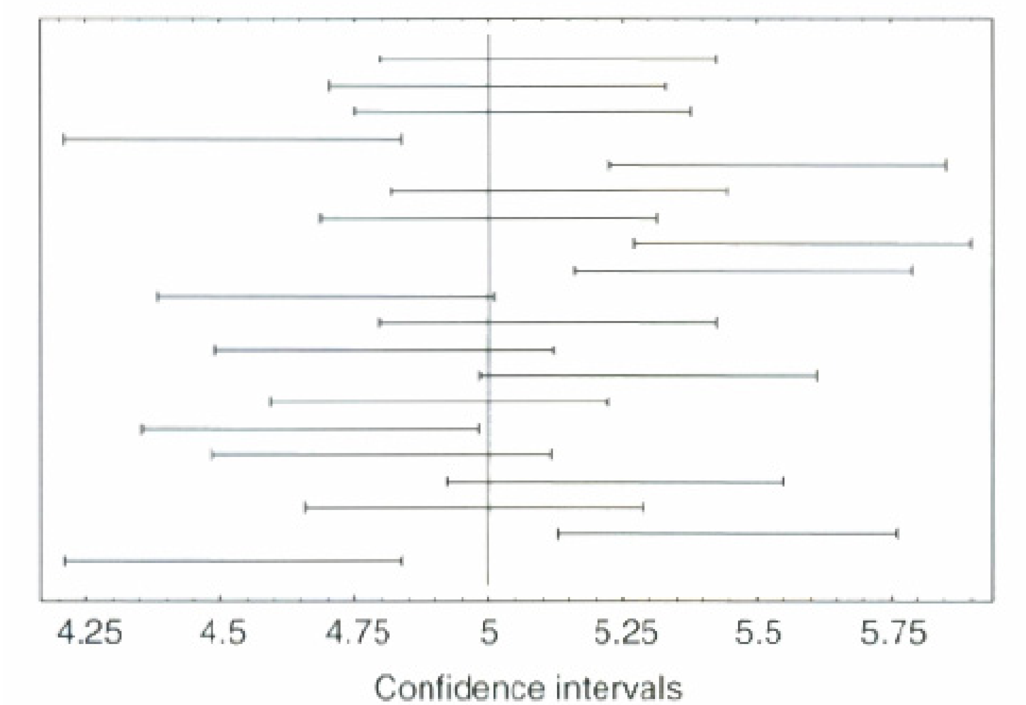
The statement $\text{Prob}\left(\bar{X} - 0.32 < \mu < \bar{X} + 0.32\right) = 0.68$

means that, if we were to repeatedly draw a large number of samples of size $n = 10$ from $N\left(\mu, \sigma^2\right)$, we expect that in 68% of these samples

$$\bar{x} - 0.32 < \mu < \bar{x} + 0.32$$

68% is known as the coverage

20 realisations of 68% confidence interval



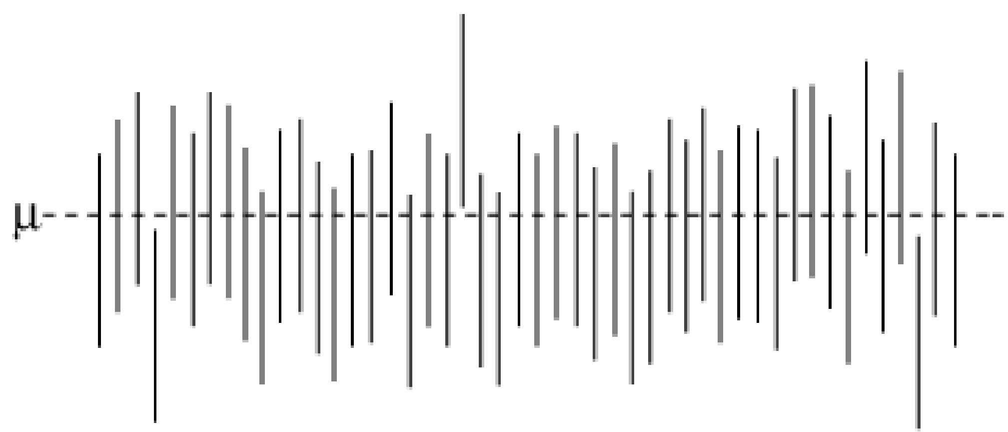
The statement $\quad \text{Prob}\left(\overline{X} - 0.32 < \mu < \overline{X} + 0.32\right) = 0.68$

means that, if we were to repeatedly draw a large number of samples of size $n = 10$ from $N\left(\mu, \sigma^2\right)$, we expect that in 68% of these samples

$$\overline{x} - 0.32 < \mu < \overline{x} + 0.32$$

68% is known as the coverage

50 realisations of 95% confidence interval



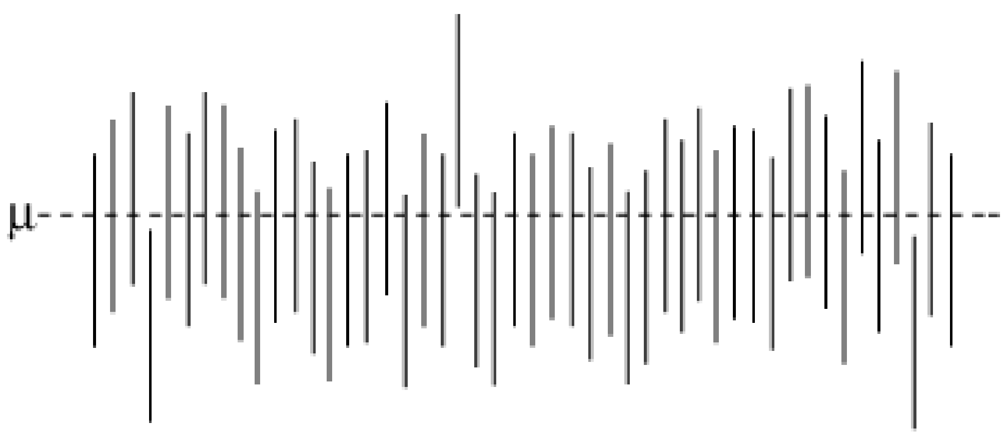
The statement $$\text{Prob}\left(\bar{X} - 0.32 < \mu < \bar{X} + 0.32\right) = 0.68$$

means that, if we were to repeatedly draw a large number of samples of size $n = 10$ from $N\left(\mu, \sigma^2\right)$, we expect that in 68% of these samples

$$\bar{x} - 0.32 < \mu < \bar{x} + 0.32$$

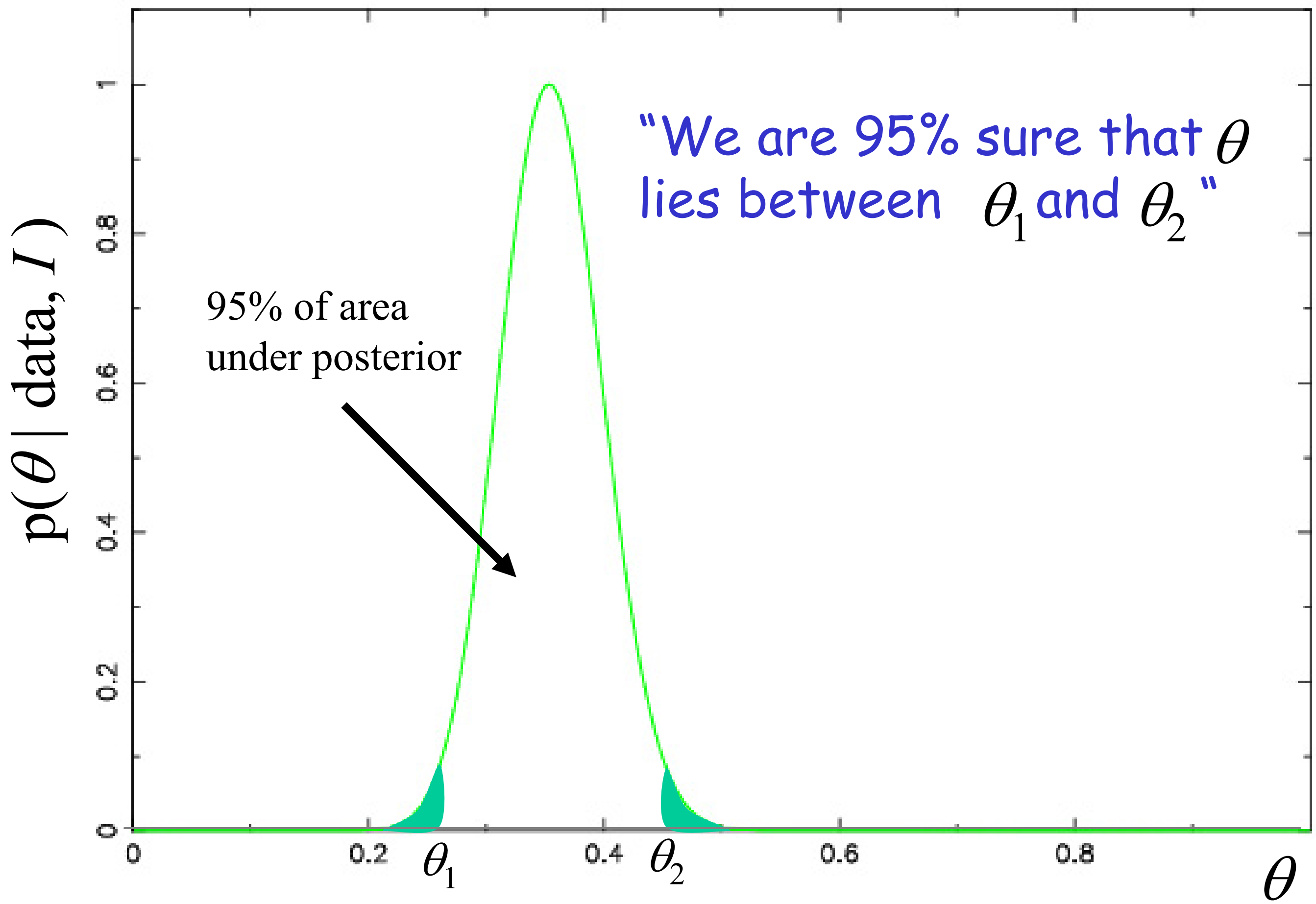
68% is known as the coverage

50 realisations of 95% confidence interval

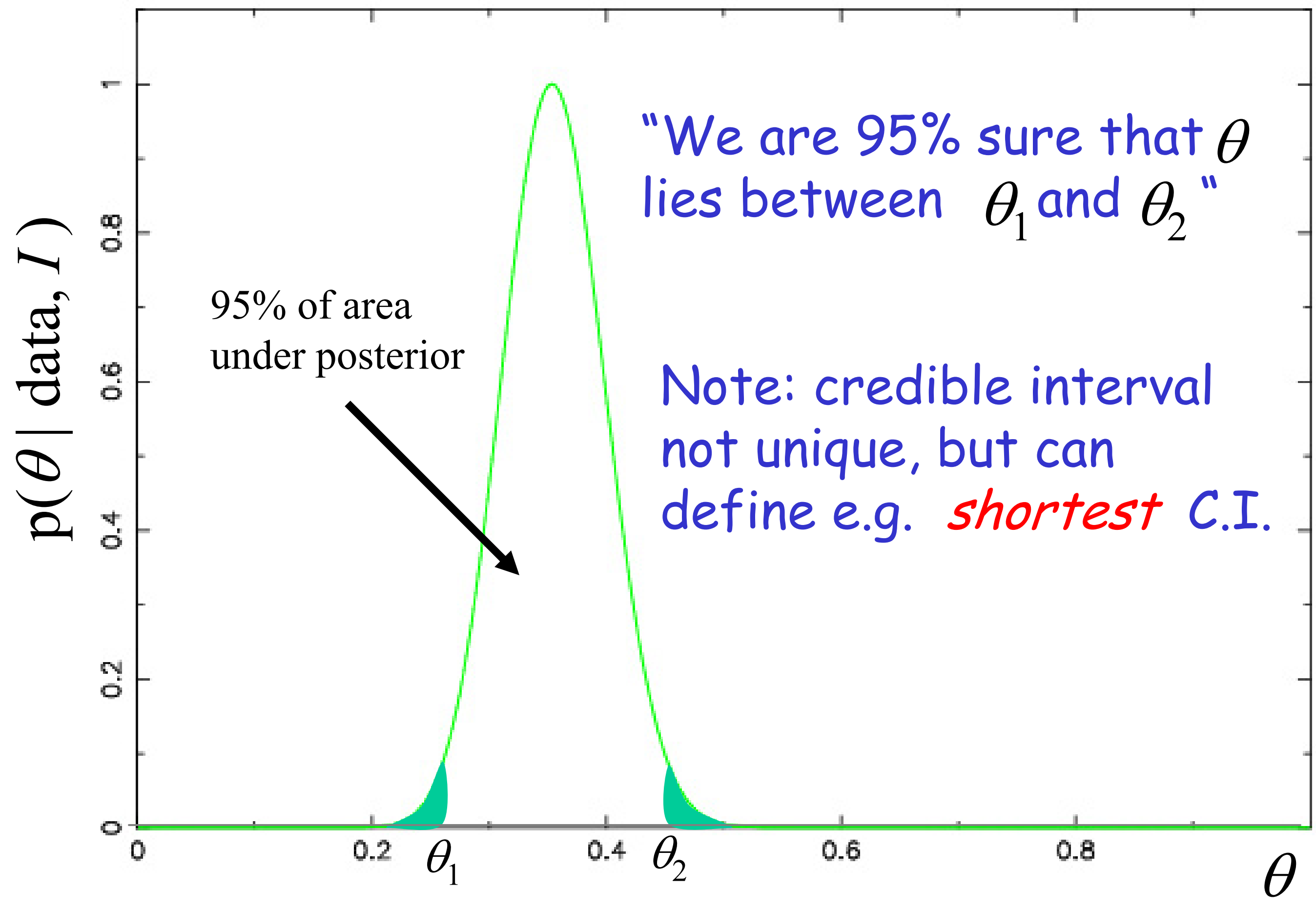


See also Mathworld demonstration

Compare the frequentist construction with Bayesian credible intervals



Compare the frequentist construction with Bayesian credible intervals



Example: Gregory, Section 14

Inference of a Poisson sampling rate

In many physics experiments the data = discrete events distributed in space, time, energy, frequency etc.

Macroscopic events: rate of earthquakes, sky location of a star

Microscopic events: LHC interactions, DM particle detections…

Model using Poisson distribution:

$$\boxed{p(n \mid r, I) = \frac{(rT)^n e^{-rT}}{n!}}$$

Model using Poisson distribution:

$$p(n \mid r, I) = \frac{(rT)^n e^{-rT}}{n!}$$

$p(n \mid r, I)$ is the probability that n discrete events will occur in time interval T, given a positive, real-valued Poisson process with event rate r, and given other background information I.

Suppose we make a single measurement of n events. From Bayes' theorem:

$$p(r \mid n, I) = \frac{p(r \mid I)\, p(n \mid r, I)}{p(n \mid I)}$$

What should we choose as our prior $p(r \mid I)$?

Later we will discuss this in more detail, and introduce the Jeffreys prior appropriate for a scale parameter.

However, the motivation for choosing a Jeffreys prior breaks down if the event rate r could be zero.

Adopt instead a uniform prior $$p(r \mid I) = \frac{1}{r_{\max}}, \quad 0 \le r \le r_{\max}$$

(See Gregory, p 377 for further discussion)

Substituting

$$p(r \mid n, I) = \Delta \frac{(rT)^n e^{-rT}}{n!}, \quad 0 \le r \le r_{\max}$$

Normalisation constant (doesn't depend on event rate)

Can show that, if $r_{\max} T >> n$ then the posterior is approximately:

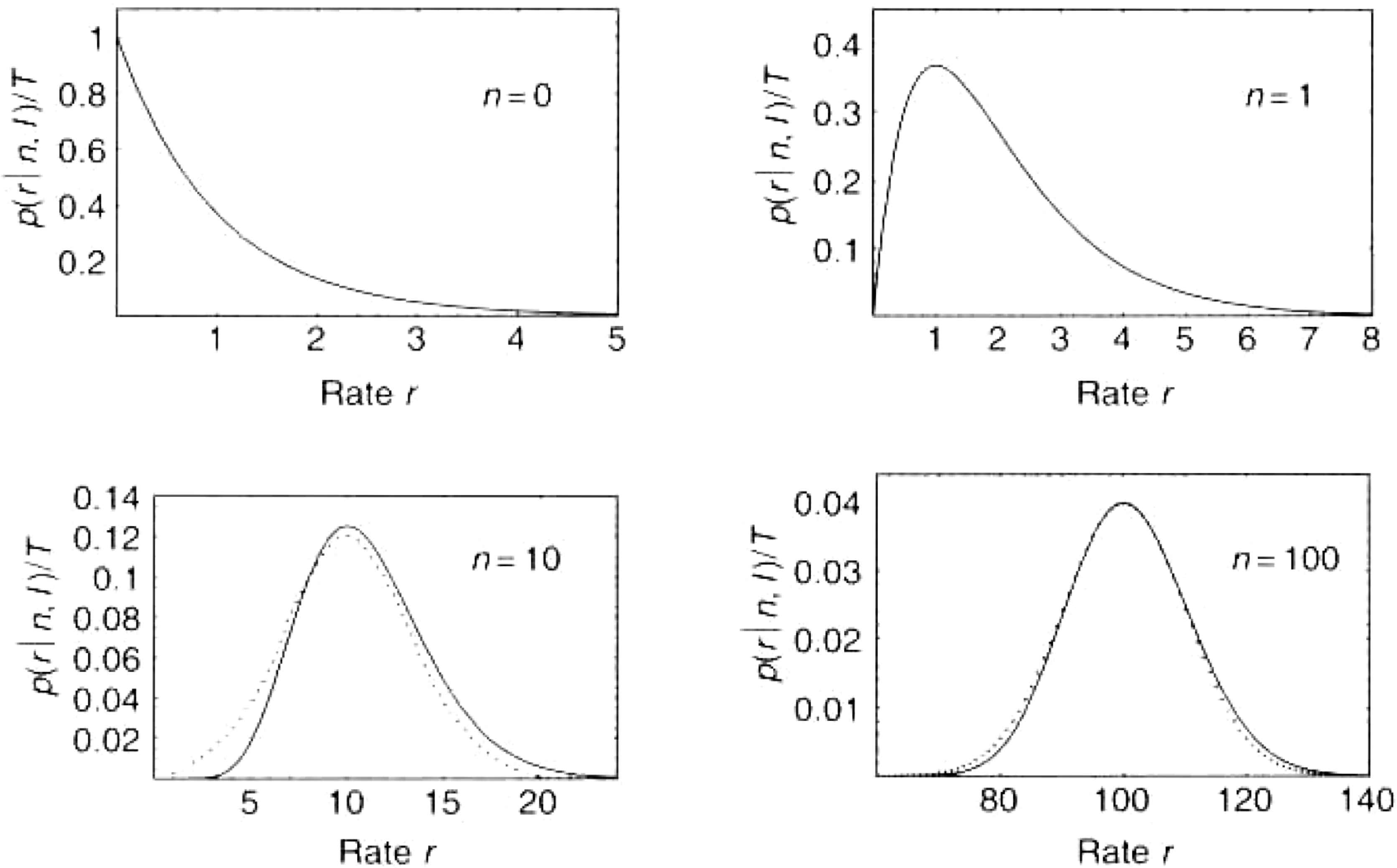
$$p(r \mid n, I) = \frac{T\,(rT)^n e^{-rT}}{n!}, \quad r \ge 0$$

$$p(r \mid n, I) = \frac{T\left(rT\right)^n e^{-rT}}{n!}, \quad r \geq 0$$

Mode: $$r_{\mathrm{mode}} = n / T$$

Mean: $$\langle r \rangle = \left(n+1\right)/T$$

Sigma $$\sigma_r = \sqrt{\left(n+1\right)} / T$$



From Gregory, pg 379

Now suppose the measured rate consists of two components:

1. A signal, of unknown rate s
2. A background, of known rate b

$$\left.\begin{array}{l} s \\ b \end{array}\right\} \quad r = s + b$$

Because we are assuming the background rate is known it follows that

$$p(s \mid n, b, I) = p(r \mid n, b, I)$$

and

$$\boxed{p(s \mid n, b, I) = C \, \frac{T\left[(s+b)T\right]^n e^{-(s+b)T}}{n!}, \quad s \geq 0}$$

Normalisation constant

$$p(s \mid n, b, I) = C \frac{T\left[(s+b)T\right]^n e^{-(s+b)T}}{n!}, \quad s \geq 0$$

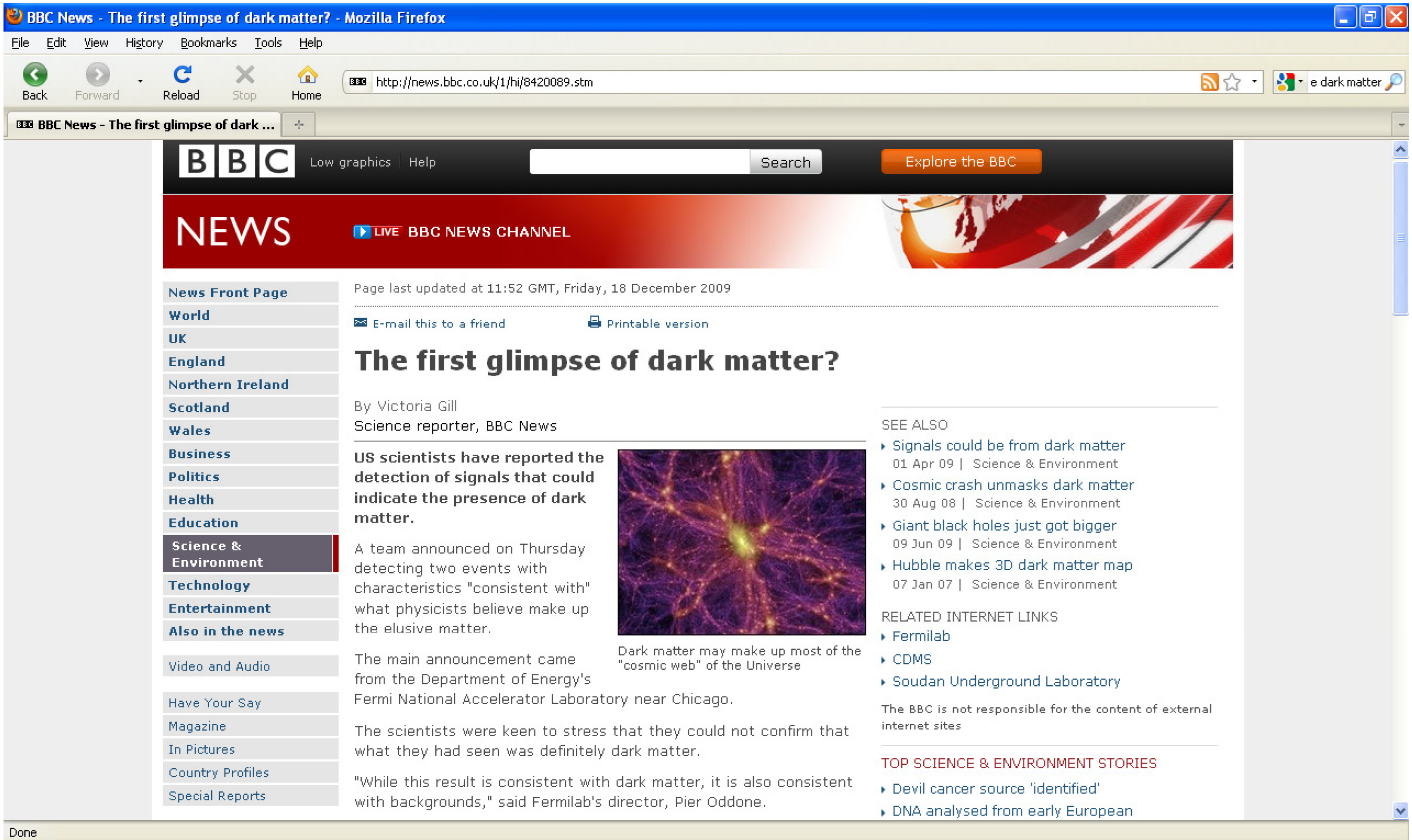
Normalisation constant

Can show that
$$C^{-1} = \sum_{i=0}^{n} \frac{(bT)^i e^{-bT}}{i!}$$

Example: Dark Matter experimental results, reported Dec 2009

Simple analysis: $n = 2$

$$b = 0.8$$



BBC News - The first glimpse of dark matter? - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Back Forward Reload Stop Home http://news.bbc.co.uk/1/hi/8420089.stm e dark matter

BBC News - The first glimpse of dark ...

BBC Low graphics Help Search Explore the BBC

NEWS LIVE BBC NEWS CHANNEL

- News Front Page
- World
- UK
- England
- Northern Ireland
- Scotland
- Wales
- Business
- Politics
- Health
- Education
- Science & Environment
- Technology
- Entertainment
- Also in the news
- Video and Audio
- Have Your Say
- Magazine
- In Pictures
- Country Profiles
- Special Reports

Page last updated at 11:52 GMT, Friday, 18 December 2009

E-mail this to a friend   Printable version

The first glimpse of dark matter?

By Victoria Gill
Science reporter, BBC News

US scientists have reported the detection of signals that could indicate the presence of dark matter.

Dark matter may make up most of the "cosmic web" of the Universe

A team announced on Thursday detecting two events with characteristics "consistent with" what physicists believe make up the elusive matter.

The main announcement came from the Department of Energy's Fermi National Accelerator Laboratory near Chicago.

The scientists were keen to stress that they could not confirm that what they had seen was definitely dark matter.

"While this result is consistent with dark matter, it is also consistent with backgrounds," said Fermilab's director, Pier Oddone.

SEE ALSO
- Signals could be from dark matter
  01 Apr 09 | Science & Environment
- Cosmic crash unmasks dark matter
  30 Aug 08 | Science & Environment
- Giant black holes just got bigger
  09 Jun 09 | Science & Environment
- Hubble makes 3D dark matter map
  07 Jan 07 | Science & Environment

RELATED INTERNET LINKS
- Fermilab
- CDMS
- Soudan Underground Laboratory

The BBC is not responsible for the content of external internet sites

TOP SCIENCE & ENVIRONMENT STORIES
- Devil cancer source 'identified'
- DNA analysed from early European

Done

Results from the Final Exposure of the CDMS II Experiment

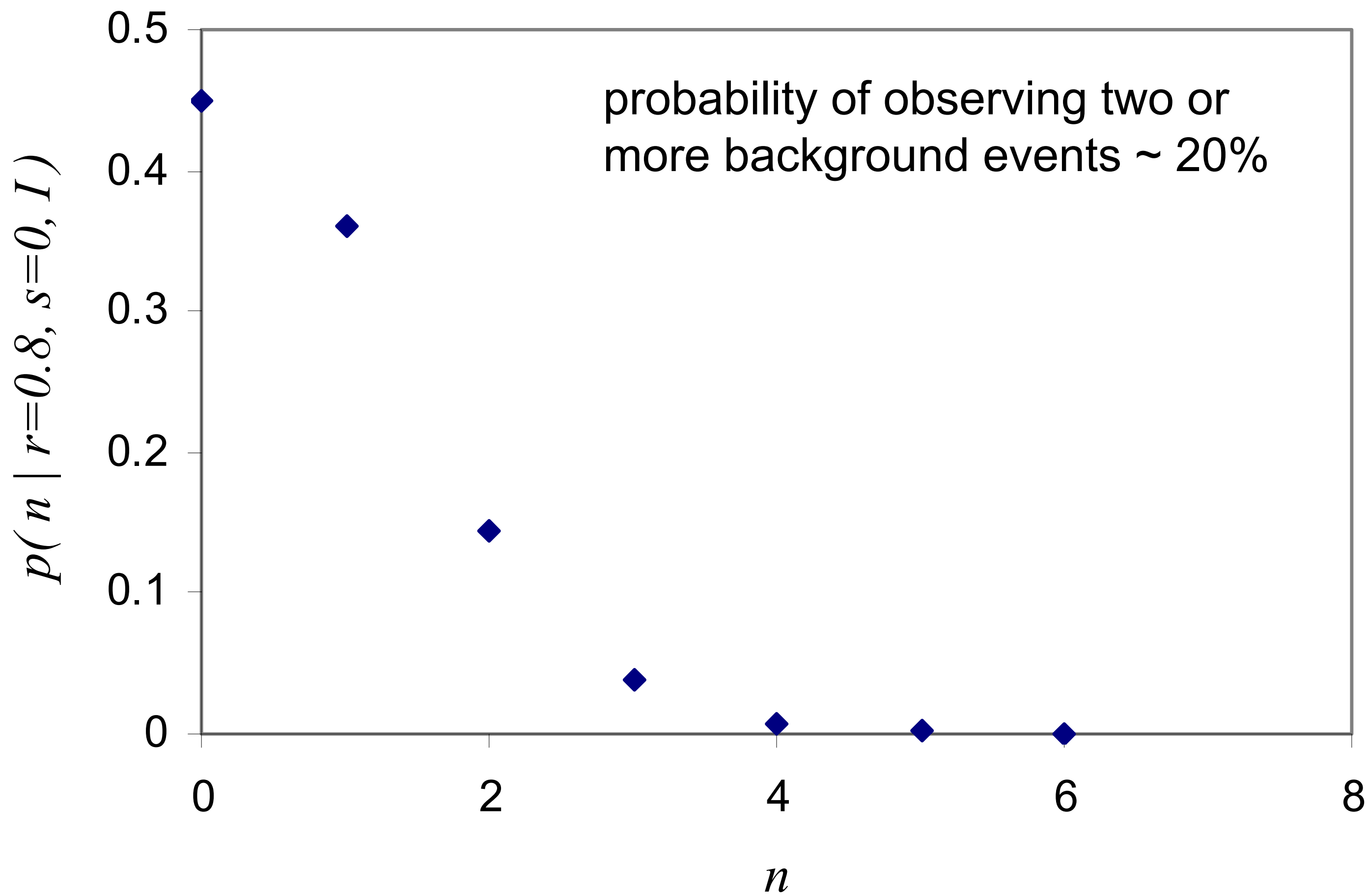
Z. Ahmed,[19] D.S. Akerib,[2] S. Arrenberg,[18] C.N. Bailey,[2] D. Balakishiyeva,[16] L. Baudis,[18] D.A. Bauer,[3] P.L. Brink,[10] T. Bruch,[18] R. Bunker,[14] B. Cabrera,[10] D.O. Caldwell,[14] J. Cooley,[9] P. Cushman,[17] M. Daal,[13] F. DeJongh,[3] M.R. Dragowsky,[2] L. Duong,[17] S. Fallows,[17] E. Figueroa-Feliciano,[5] J. Filippini,[19] M. Fritts,[17] S.R. Golwala,[19] D.R. Grant,[2] J. Hall,[3] R. Hennings-Yeomans,[2] S.A. Hertel,[5] D. Holmgren,[3] L. Hsu,[3] M.E. Huber,[15] O. Kamaev,[17] M. Kiveni,[11] M. Kos,[11] S.W. Leman,[5] R. Mahapatra,[12] V. Mandic,[17] K.A. McCarthy,[5] N. Mirabolfathi,[13] D. Moore,[19] H. Nelson,[14] R.W. Ogburn,[10] A. Phipps,[13] M. Pyle,[10] X. Qiu,[17] E. Ramberg,[3] W. Rau,[6] A. Reisetter,[17,7] T. Saab,[16] B. Sadoulet,[4,13] J. Sander,[14] R.W. Schnee,[11] D.N. Seitz,[13] B. Serfass,[13] K.M. Sundqvist,[13] M. Tarka,[18] P. Wikus,[5] S. Yellin,[10,14] J. Yoo,[3] B.A. Young,[8] and J. Zhang[17]

(CDMS Collaboration)

[1] *Division of Physics, Mathematics & Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*
[2] *Department of Physics, Case Western Reserve University, Cleveland, OH 44106, USA*
[3] *Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*
[4] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[5] *Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[6] *Department of Physics, Queen's University, Kingston, ON, Canada, K7L 3N6*
[7] *Department of Physics, St. Olaf College, Northfield, MN 55057 USA*
[8] *Department of Physics, Santa Clara University, Santa Clara, CA 95053, USA*
[9] *Department of Physics, Southern Methodist University, Dallas, TX 75275, USA*
[10] *Department of Physics, Stanford University, Stanford, CA 94305, USA*
[11] *Department of Physics, Syracuse University, Syracuse, NY 13244, USA*
[12] *Department of Physics, Texas A & M University, College Station, TX 77843, USA*
[13] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[14] *Department of Physics, University of California, Santa Barbara, CA 93106, USA*
[15] *Departments of Phys. & Elec. Engr., University of Colorado Denver, Denver, CO 80217, USA*
[16] *Department of Physics, University of Florida, Gainesville, FL 32611, USA*
[17] *School of Physics & Astronomy, University of Minnesota, Minneapolis, MN 55455, USA*
[18] *Physics Institute, University of Zürich, Winterthurerstr. 190, CH-8057, Switzerland*
[19] *Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*

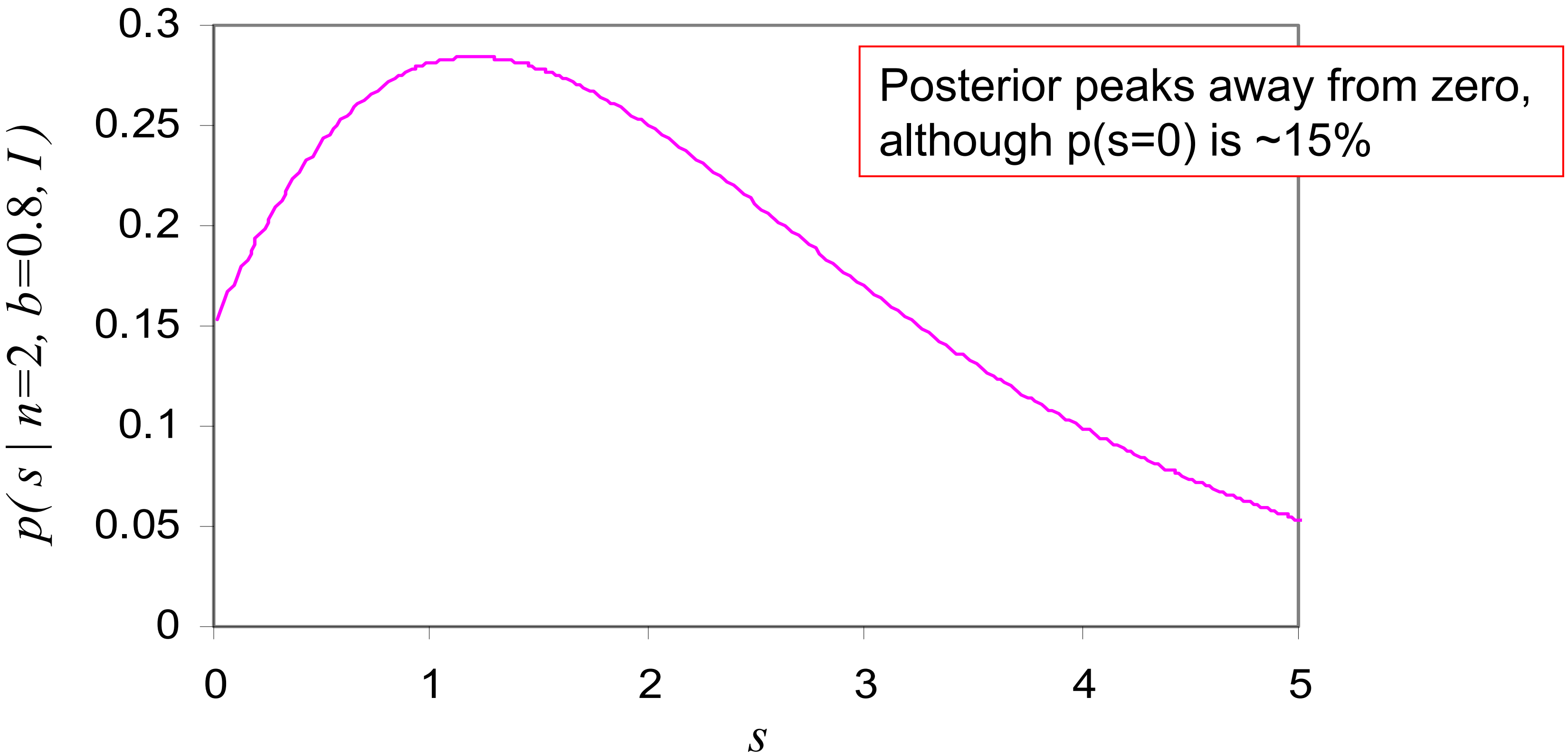
We report results from a blind analysis of the final data taken with the Cryogenic Dark Matter Search experiment (CDMS II) at the Soudan Underground Laboratory, Minnesota, USA. A total raw exposure of 612 kg-days was analyzed for this work. We observed two events in the signal region; based on our background estimate, the probability of observing two or more background events is 23%. These data set an upper limit on the Weakly Interacting Massive Particle (WIMP)-nucleon elastic-scattering spin-independent cross-section of 7.0×10^{-44} cm^2 for a WIMP of mass 70 GeV/c^2 at the 90% confidence level. Combining this result with all previous CDMS II data gives an upper limit on the WIMP-nucleon spin-independent cross-section of 3.8×10^{-44} cm^2 for a WIMP of mass 70 GeV/c^2. We also exclude new parameter space in recently proposed inelastic dark matter models.

Predicted event rate, assuming **no** signal $p(n \mid r=0.8, I) = \frac{(0.8\,T)^n e^{-0.8\,T}}{n!}$



Posterior pdf for the signal rate

$$p(s \mid n=2, b=0.8, I) = C\,\frac{\left[(s+0.8)\right]^2 e^{-(s+0.8)}}{2!}, \quad s \geq 0$$



Results from the Final Exposure of the CDMS II Experiment

Z. Ahmed,[19] D.S. Akerib,[2] S. Arrenberg,[18] C.N. Bailey,[2] D. Balakishiyeva,[16] L. Baudis,[18] D.A. Bauer,[3] P.L. Brink,[10] T. Bruch,[18] R. Bunker,[14] B. Cabrera,[10] D.O. Caldwell,[14] J. Cooley,[9] P. Cushman,[17] M. Daal,[13] F. DeJongh,[3] M.R. Dragowsky,[2] L. Duong,[17] S. Fallows,[17] E. Figueroa-Feliciano,[5] J. Filippini,[19] M. Fritts,[17] S.R. Golwala,[19] D.R. Grant,[2] J. Hall,[3] R. Hennings-Yeomans,[2] S.A. Hertel,[5] D. Holmgren,[3] L. Hsu,[3] M.E. Huber,[15] O. Kamaev,[17] M. Kiveni,[11] M. Kos,[11] S.W. Leman,[5] R. Mahapatra,[12] V. Mandic,[17] K.A. McCarthy,[5] N. Mirabolfathi,[13] D. Moore,[19] H. Nelson,[14] R.W. Ogburn,[10] A. Phipps,[13] M. Pyle,[10] X. Qiu,[17] E. Ramberg,[3] W. Rau,[6] A. Reisetter,[17,7] T. Saab,[16] B. Sadoulet,[4,13] J. Sander,[14] R.W. Schnee,[11] D.N. Seitz,[13] B. Serfass,[13] K.M. Sundqvist,[13] M. Tarka,[18] P. Wikus,[5] S. Yellin,[10,14] J. Yoo,[3] B.A. Young,[8] and J. Zhang[17]

(CDMS Collaboration)

[1] *Division of Physics, Mathematics & Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*
[2] *Department of Physics, Case Western Reserve University, Cleveland, OH 44106, USA*
[3] *Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*
[4] *Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*
[5] *Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[6] *Department of Physics, Queen's University, Kingston, ON, Canada, K7L 3N6*
[7] *Department of Physics, St. Olaf College, Northfield, MN 55057 USA*
[8] *Department of Physics, Santa Clara University, Santa Clara, CA 95053, USA*
[9] *Department of Physics, Southern Methodist University, Dallas, TX 75275, USA*
[10] *Department of Physics, Stanford University, Stanford, CA 94305, USA*
[11] *Department of Physics, Syracuse University, Syracuse, NY 13244, USA*
[12] *Department of Physics, Texas A & M University, College Station, TX 77843, USA*
[13] *Department of Physics, University of California, Berkeley, CA 94720, USA*
[14] *Department of Physics, University of California, Santa Barbara, CA 93106, USA*
[15] *Departments of Phys. & Elec. Engr., University of Colorado Denver, Denver, CO 80217, USA*
[16] *Department of Physics, University of Florida, Gainesville, FL 32611, USA*
[17] *School of Physics & Astronomy, University of Minnesota, Minneapolis, MN 55455, USA*
[18] *Physics Institute, University of Zürich, Winterthurerstr. 190, CH-8057, Switzerland*
[19] *Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*

912.3592v1 [astro-ph.CO] 18 Dec 2009

We report results from a blind analysis of the final data taken with the Cryogenic Dark Matter Search experiment (CDMS II) at the Soudan Underground Laboratory, Minnesota, USA. A total raw exposure of 612 kg-days was analyzed for this work. We observed two events in the signal region; based on our background estimate, the probability of observing two or more background events is 23%. These data set an upper limit on the Weakly Interacting Massive Particle (WIMP)-nucleon elastic-scattering spin-independent cross-section of 7.0×10^{-44} cm^2 for a WIMP of mass 70 GeV/c^2 at the 90% confidence level. Combining this result with all previous CDMS II data gives an upper limit on the WIMP-nucleon spin-independent cross-section of 3.8×10^{-44} cm^2 for a WIMP of mass 70 GeV/c^2. We also exclude new parameter space in recently proposed inelastic dark matter models.

University of Glasgow

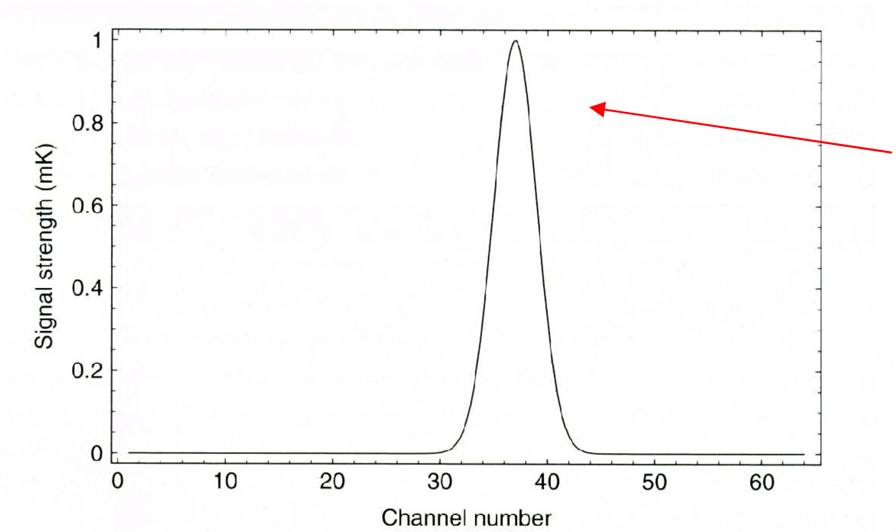
Further example: Gregory, Section 3.6

Fitting the amplitude of a spectral line.

Model M1: Signal strength $= T \exp\left\{ \frac{-(\nu_i - \nu_o)^2}{2\sigma_L^2} \right\}$

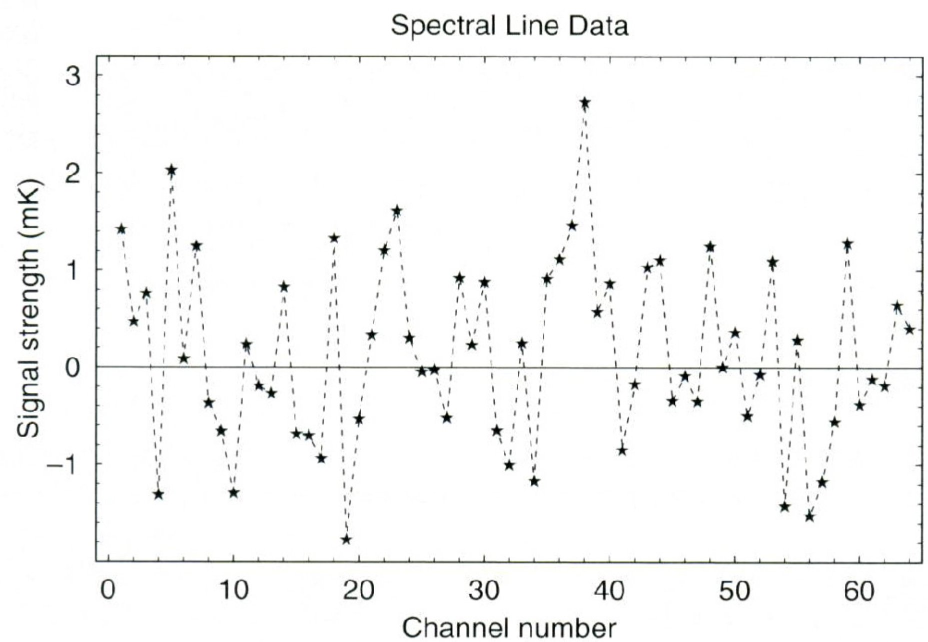
Amplitude

Assume other parameters are known



Parameter to be fitted is amplitude of signal (taken here to be unity)

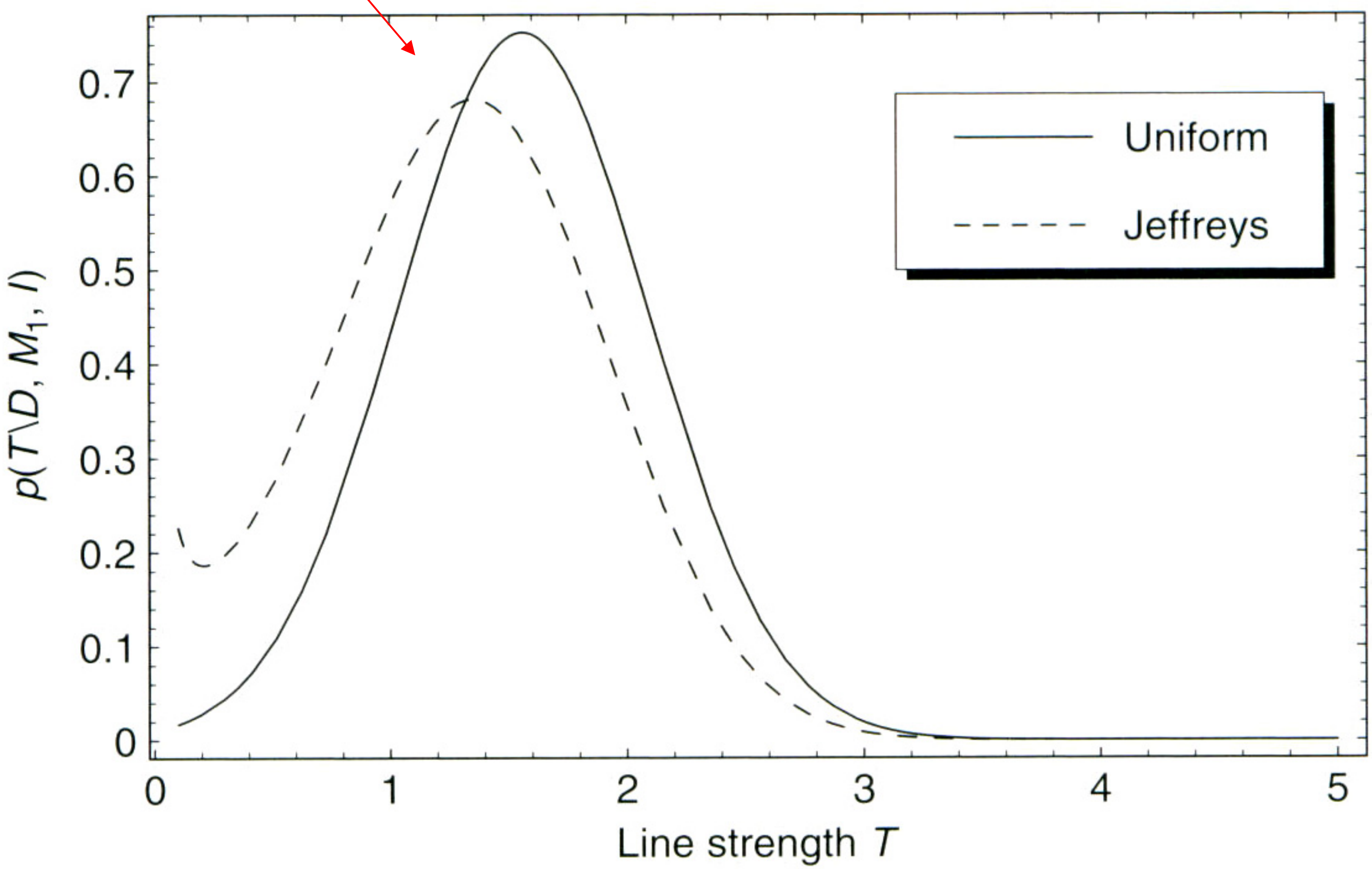
Observed data

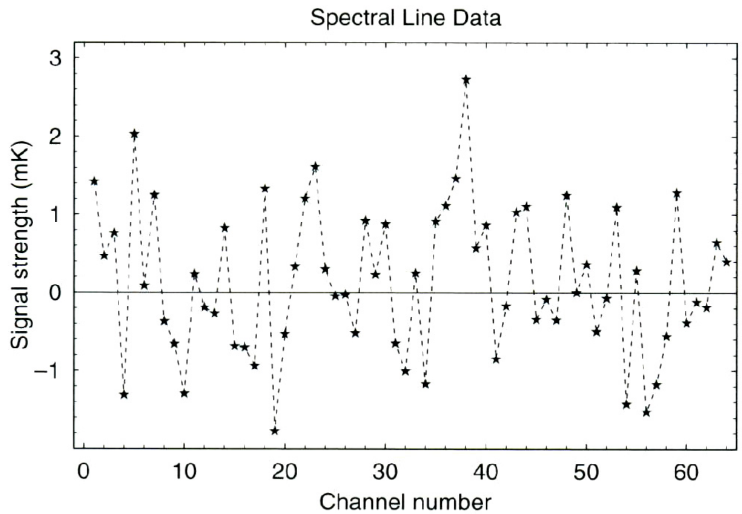


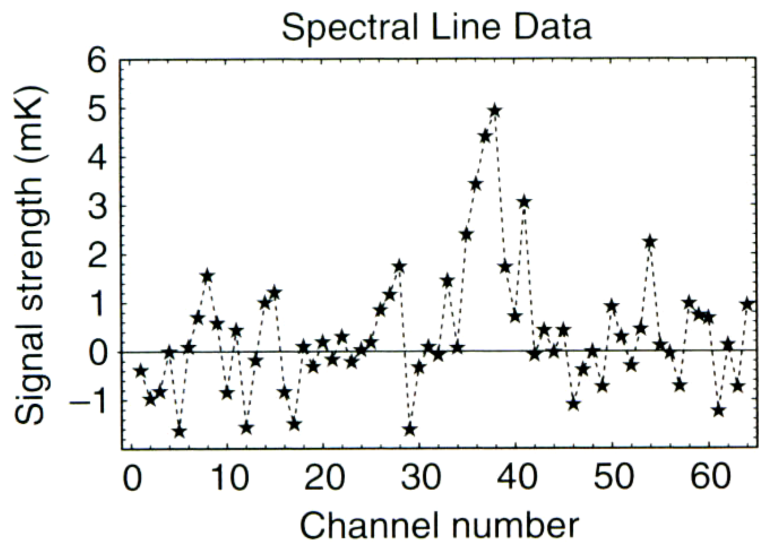
$$p(T \mid D, M_1, I) \propto p(T \mid M_1, I)\, p(D \mid M_1, T, I)$$

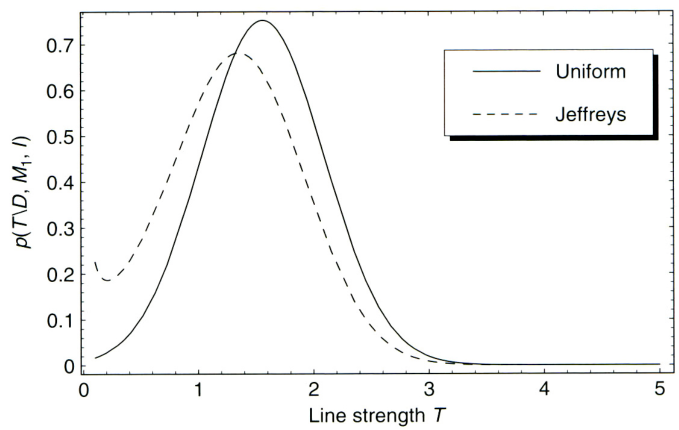
posterior prior likelihood

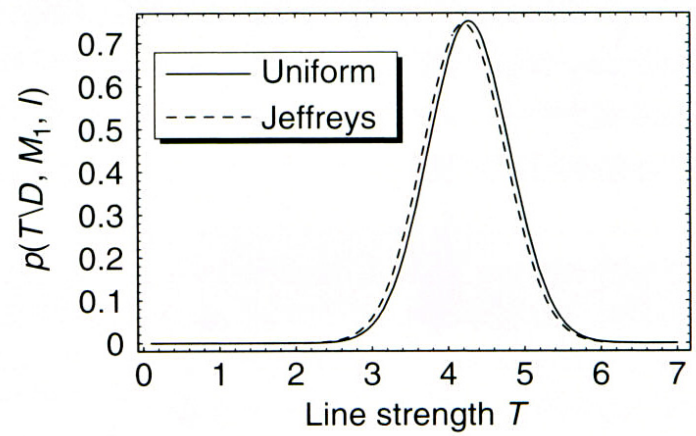
Posterior sensitive to choice of prior (see later)







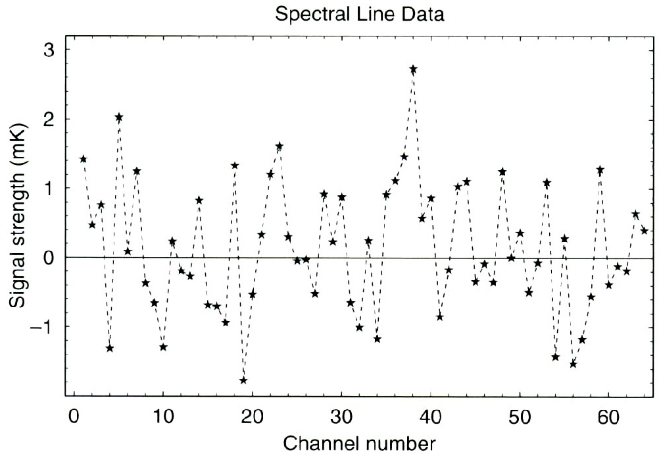




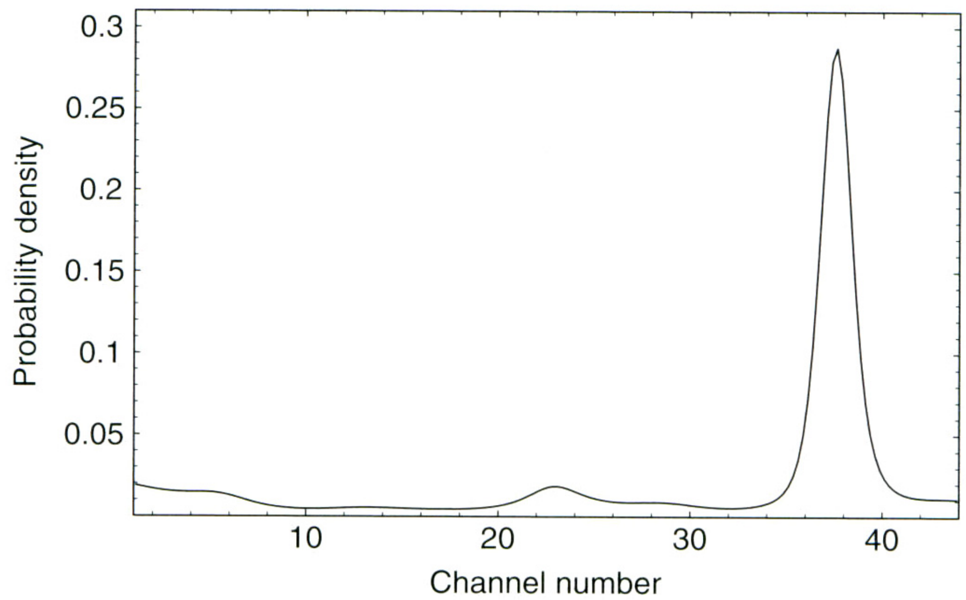
Prior dependence less strong for stronger signal

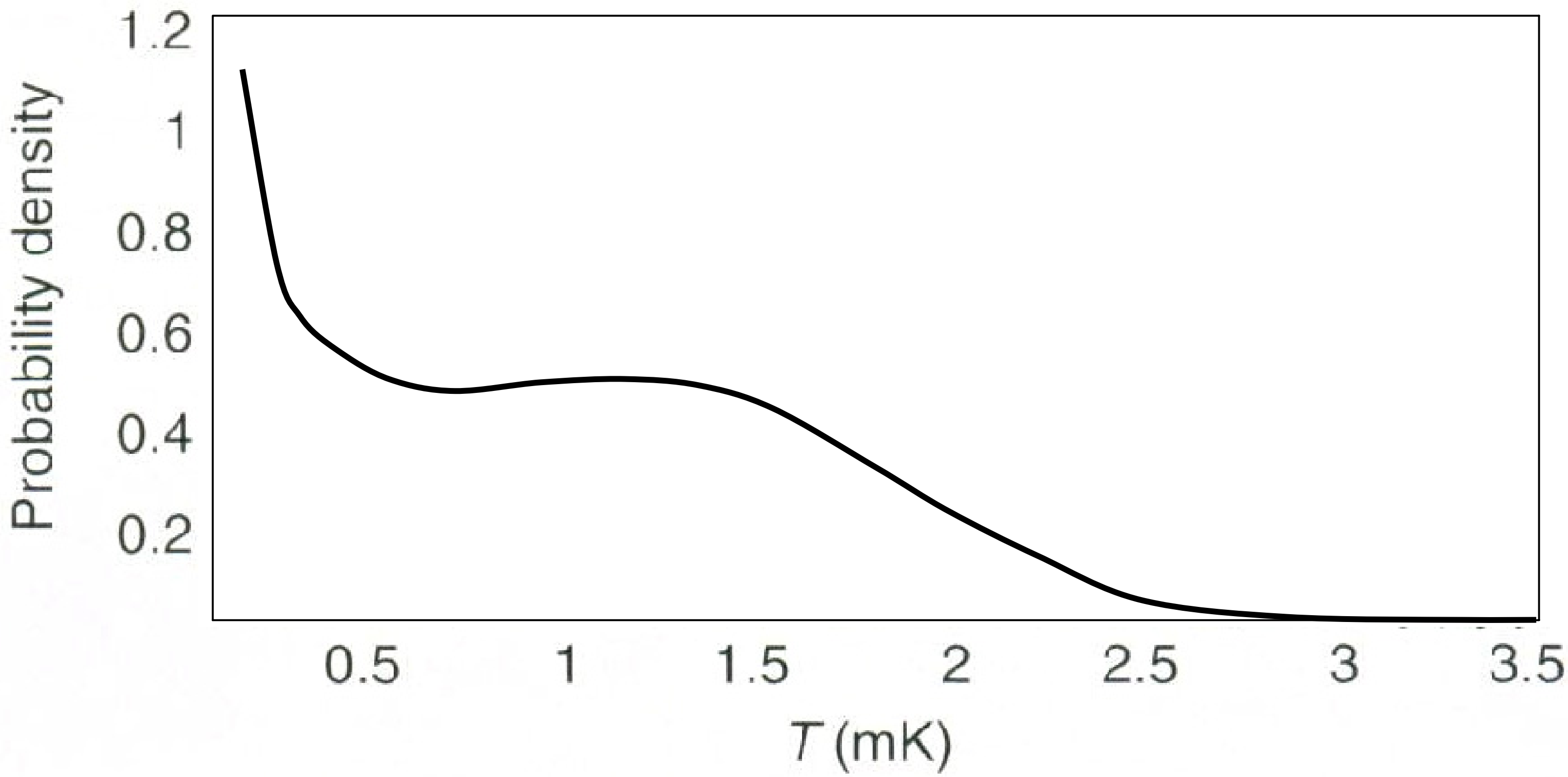
What if the frequency (channel number) is also unknown?

Can compute marginal posterior for C.N.



Possible lines at ~20 and < 10





Allowing the C.N. to be a free parameter changes significantly the marginal posterior for the amplitude.

(But should we be fitting only one line? See Section 5)

Bayesian versus Frequentist statistics: Who is right?

If we adopt a uniform prior, results of Bayesian estimation are formally equivalent to maximum likelihood

Posterior Likelihood Prior

$$\boxed{\underbrace{p(\text{model} \mid \text{data}, I)}_{\text{Posterior}} \propto \underbrace{p(\text{data} \mid \text{model}, I)}_{\text{Likelihood}} \times \underbrace{p(\text{model} \mid I)}_{\text{Prior}}}$$

But underlying principle is completely different.

(and often we should *not* assume a uniform prior – see later)

Bayesian versus Frequentist statistics: Who is right?

If we adopt a uniform prior, results of Bayesian estimation are formally equivalent to maximum likelihood

$$\underbrace{p(\text{model} \mid \text{data}, I)}_{\text{Posterior}} \propto \underbrace{p(\text{data} \mid \text{model}, I)}_{\text{Likelihood}} \times \underbrace{p(\text{model} \mid I)}_{\text{Prior}}$$

But underlying principle is completely different.

(and often we should *not* assume a uniform prior – see later)

"Fundamentalist" views expressed on both sides:

See my.SUPA site for some references.

Bayesian Analysis (2008) **3**, Number 3, pp. 445–450

Objections to Bayesian statistics

Andrew Gelman*

Abstract. Bayesian inference is one of the more controversial approaches to statistics. The fundamental objections to Bayesian methods are twofold: on one hand, Bayesian methods are presented as an automatic inference engine, and this raises suspicion in anyone with applied experience. The second objection to Bayes comes from the opposite direction and addresses the subjective strand of Bayesian inference. This article presents a series of objections to Bayesian inference, written in the voice of a hypothetical anti-Bayesian statistician. The article is intended to elicit elaborations and extensions of these and other arguments from non-Bayesians and responses from Bayesians who might have different perspectives on these issues.

Keywords: Foundations, Comparisons to other methods

1 A Bayesian's attempt to see the other side

Bayesian inference is one of the more controversial approaches to statistics, with both the promise and limitations of being a closed system of logic. There is an extensive literature, which sometimes seems to overwhelm that of Bayesian inference itself, on the advantages and disadvantages of Bayesian approaches. Bayesians' contributions to this discussion have included defense (explaining how our methods reduce to classical methods as special cases, so that we can be as inoffensive as anybody if needed), affirmation (listing the problems that we can solve more effectively as Bayesians), and attack (pointing out gaps in classical methods).

The present article is unusual in representing a Bayesian's presentation of what he views as the strongest non-Bayesian arguments. Although this originated as an April Fool's blog entry (Gelman, 2008), I realized that these are strong arguments to be taken seriously—and ultimately accepted in some settings and refuted in others.

Bayesian Analysis (2008) 0, Number 0, pp. 1–6

This Physicist's view of Gelman's Bayes

John Skilling
Maximum Entropy Data Consultants Ltd.
Killaha East, Kenmare, County Kerry, Ireland
skilling@eircom.net — August 2008

Abstract. The author offers fundamentalist commentary on Andrew Gelman's brilliantly provocative comments on Bayes, and the associated discussion.

Keywords: evidence, fundamentals, semi-Bayesian, Emperor's Clothes.

1 Introduction

In publishing Gelman (2008) with commentaries, the Editor is to be congratulated on allowing an exhilirating relaxation of the orthodox norms of professional presentation. Consequently, each author's contribution is seen with unusual clarity. There's no fussy detail. There's no intricate symbolism designed to impress. There is just the natural language of personal communication, so well suited to discussion of basic outlooks.

Yet that very clarity exposes what is oddly missing. The discussions lack any serious account of why we MUST use Bayes or of how I think we SHOULD use Bayes. Readers would think there was a choice. There isn't. Here in complimentary response to Andrew's wonderfully successful provocation is my own polemical rant on the subject.

2 Why we MUST use Bayes

Probability calculus, often called "Bayesian", is not an option to be accepted, modified or rejected at whim. It has a firm logical basis as the unique calculus of rationality. Over sixty years ago, Richard Cox wrote a remarkable paper (Cox (1946)) which Jaynes (2003) considered to be "the most important advance in the conceptual (as opposed to the purely mathematical) formulation of probability theory since Laplace". I have long concurred with that view, except that I omit the bracketed qualification. Although some of us continue to polish and refine the approach, I hold that Cox (1946) remains the foundation authority.

Bayesian versus Frequentist statistics: Who is right?

If we adopt a uniform prior, results of Bayesian estimation are formally equivalent to maximum likelihood

Posterior   Likelihood   Prior

$$\boxed{p(\text{model} \mid \text{data}, I) \propto p(\text{data} \mid \text{model}, I) \times p(\text{model} \mid I)}$$

But underlying principle is completely different.

(and often we should *not* assume a uniform prior – see later)

Important to understand both Bayesian and Frequentist approaches, and always to think carefully about their applicability to your particular problem.

Bayesian versus Frequentist statistics: Who is right?

Quote from Louis Lyons

Bayesians address the question everyone is interested in by using assumptions that no one believes.

Frequentists use impeccable logic to deal with an issue of no interest to anyone.

Louis Lyons
Academic Lecture at Fermilab
August 17, 2004