Conditional Probability $\quad P(A|B) = \dfrac{P(A,B)}{P(B)}$ $\quad$ (Def)

$P(A,B)$ → P(A and B)

$$P(A,B) = P(B,A) \implies P(A|B) = P(B|A)\,\frac{P(A)}{P(B)}$$

$A \rightarrow$ model M

$B \rightarrow$ data D

these are needed for inference.

new notation

$$\overbrace{\boxed{P(M|D)}}^{\text{Posterior}} = \frac{\overbrace{\boxed{P(D|M)}}^{\text{likelihood}}\ \overbrace{\boxed{P(M)}}^{\text{Prior}}}{\underbrace{\boxed{P(D)}}_{\text{evidence.}}}$$

likelihood $P(D|M) = \mathcal{L}(D|M)$

$= \mathcal{L}(x|\theta)$

e.g. 1 data point

single model parameter.

inputs

prior $P(M) = \pi(\theta) = \pi(M) =$

posterior $P(M|D) = P(\theta|x)$

evidence $P(D) = Z$

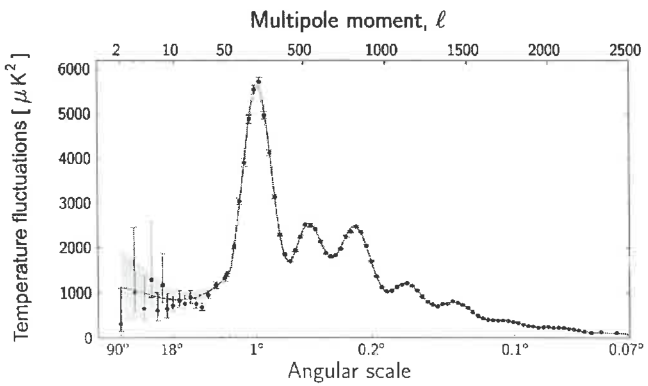
$$Z\,P(\theta|x) = \mathcal{L}(x|\theta)\,\pi(\theta)$$

The outcome of flipping a coin; H or T ?

Will it rain tomorrow ? (Is it raining in London now ?)

What is the 10^{100}th digit of π ?

Will a particular ^{14}C atom decay in the next century ?



What's the probability of finding a particular value for a specific harmonic in this CMB data ?

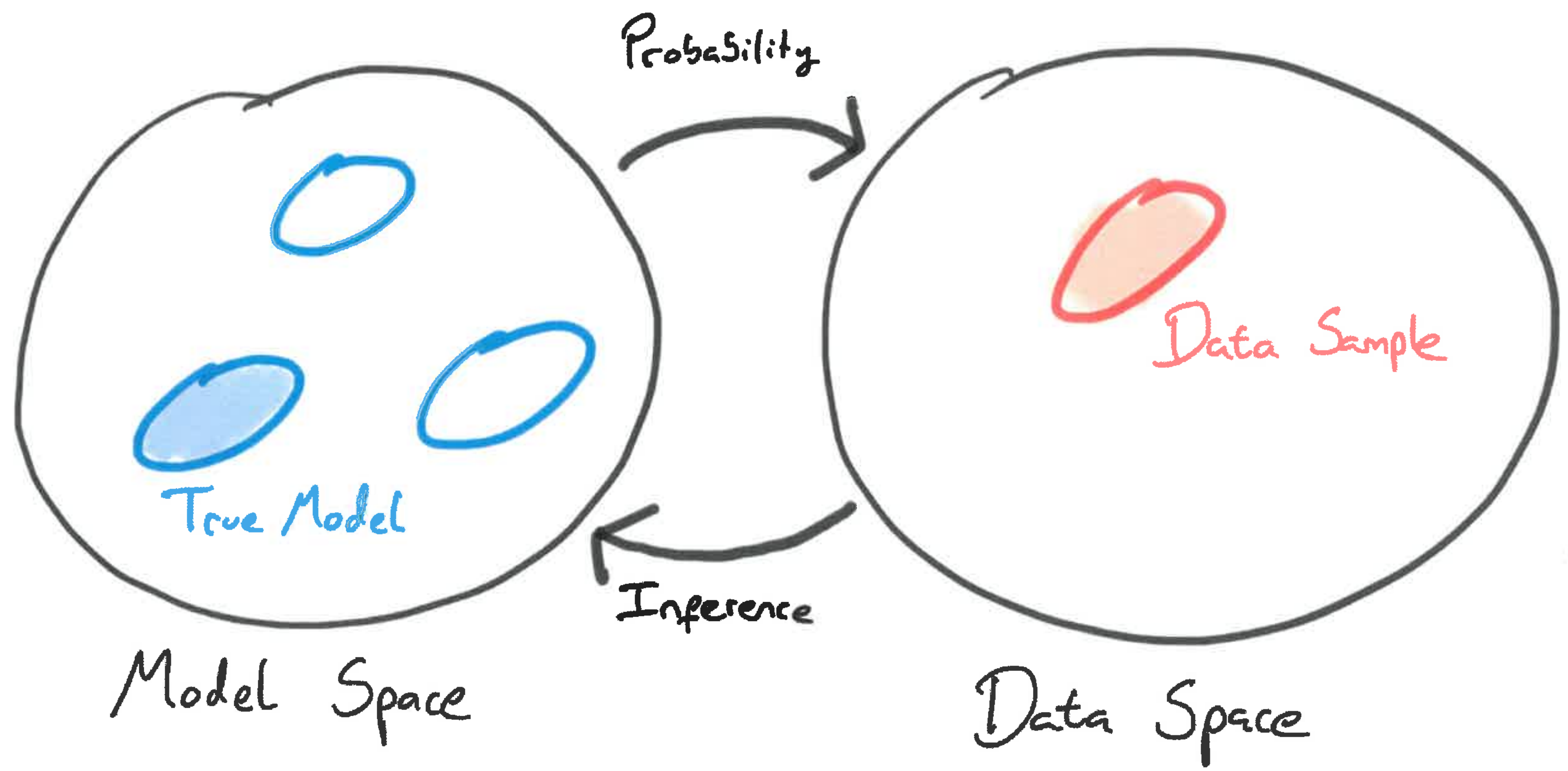
Who should win this chess game ? (Given perfect play.)

Will I roll a double 6 on my next turn in Backgammon ?

$P(s|n)$

$\rightarrow$ random

$\mathcal{L}(n|s)$



example problem (Poisson)

estimating S from survey counted n stars in area A

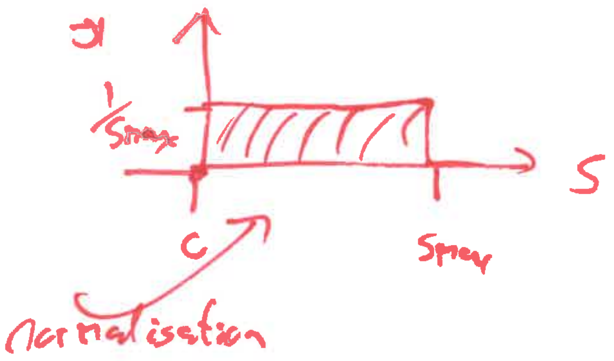
$$\mathcal{L}(n|S) = \frac{(AS)^n \exp(-AS)}{n!}$$

$$P(S|n) = \frac{\mathcal{L}(n|S)\ \pi(S)}{Z}$$

To start we need to choose $\pi(s)$.

"Ignorance prior"

$$\pi(s) = \begin{cases} 1/s_{max} & s_{max} \geq s > 0 \\ 0 & \text{else} \end{cases}$$



$$\pi(s) = \mathbb{1}_{(0,s_{max})}(s)\,\frac{1}{s_{max}}$$

indicator function

Using Bayes' theorem w uniform prior

$$P(s|n) = \frac{(As)^n e^{-As}}{Z\, n!\, s_{max}} \mathbb{1}(s)$$

evidence from normalisation $\Rightarrow$

$$Z = \int_0^{s_{max}} \frac{(As)^n e^{-As}}{n!\, s_{max}}\, ds$$

prior choices.

$$\mathcal{L}(n|S) = \frac{(AS)^n e^{-AS}}{n!}$$

- PDF on data
- function of model

$$\sum_{n=0}^{\infty} \mathcal{L}(n|S) = 1$$

$$\int_0^{\infty} dS\, \mathcal{L}(n|S) \neq 1$$

when choosing π, should reflect our knowledge.

S is scale parameter

$$\pi(\cancel{x}\,S)\, d\cancel{x}\,S = \pi(\alpha \cancel{x}\,S)\, d(\alpha \cancel{x}\,S)$$ ← want this to be true for any $\alpha > 0$.

$$\pi(\cancel{x}\,S) \propto \frac{1}{\cancel{x}\,S}$$

2 choices so far for prior

π uniform $\propto$ const

π scale invariant $\propto \frac{1}{S}$

$$\pi(s) = \frac{1}{\log\left(\frac{s_{max}}{s_{min}}\right)} \begin{cases} \frac{1}{s} & s_{min} < s < s_{max} \\ 0 & \text{else} \end{cases}$$

properly normalised, scale-invariant prior

"Jeffrey's prior"