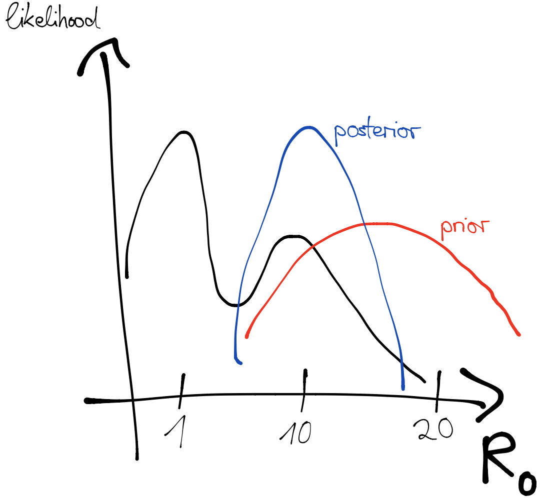
Day 4

Centre for the Mathematical Modelling of Infectious Diseases
London School of Hygiene & Tropical Medicine

The point of all of this

- posterior probabilities

$$p(\theta|\text{data}) = \frac{p(\text{data}|\theta)p(\theta)}{p(\text{data})}$$



Sampling from the posterior

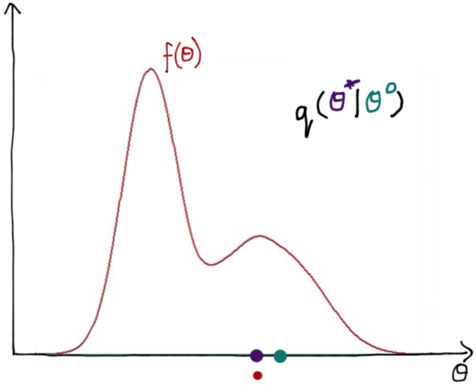
We interpret $p(\theta|\mathrm{data})$ as the probability distribution of a random variable θ, from which we ***sample*** (via MCMC)

Why sample?

1. explore parameter space
2. samples can be useful
   - explore interventions, forecasts

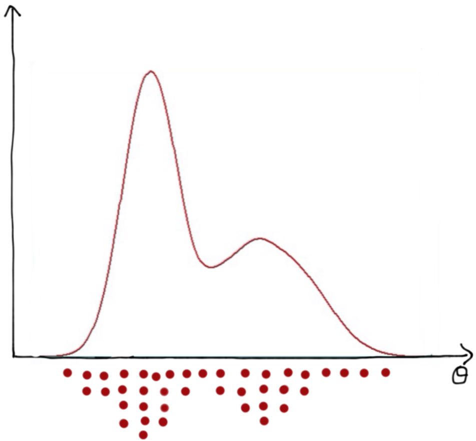
MCMC: Sampling from a distribution

- We can calculate (in a deterministic model) $p(\theta|\text{data})$ given any θ – how do we sample?



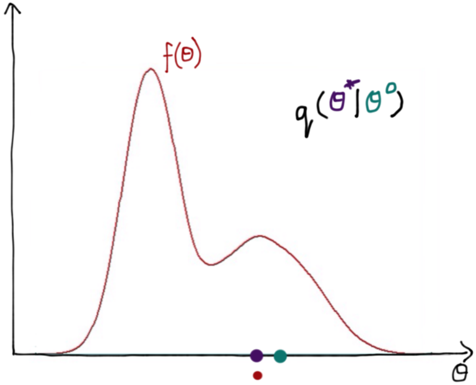
MCMC: Sampling from a distribution

- We can calculate (in a deterministic model) $p(\theta|\text{data})$ given any θ – how do we sample?



MCMC: Sampling from a distribution

- We can calculate (in a deterministic model) $p(\theta|\text{data})$ given any θ – how do we sample?



Stochastic models

- one θ can lead to many possible outcomes X
- we can
  1. sample from $p(X|\theta)$ (via simulation)
  2. evaluate the trajectory likelihood $p(\mathrm{data}|X,\theta)$
- we can't directly evaluate the likelihood $p(\mathrm{data}|\theta)$

$$p(\mathrm{data}|\theta) = \sum_X p(\mathrm{data}|X,\theta)p(X|\theta)$$

- The number of possible trajectories X for one value of θ is large, potentially infinite

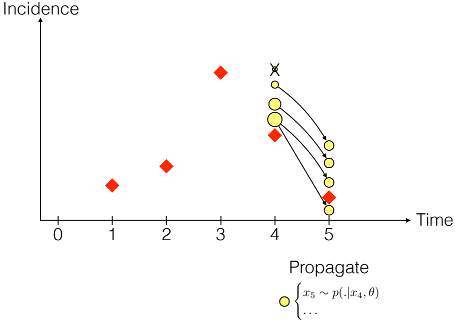
Sequential Monte Carlo (SMC) / Particle Filter I

To approximate

$$p(\mathrm{data}|\theta) = \sum_X p(\mathrm{data}|X, \theta) p(X|\theta)$$

we sample n trajectories X_n from

$$p(Y_{1:(T-1)}|X, \theta) p(X|\theta)$$



Sequential Monte Carlo (SMC) / Particle Filter II

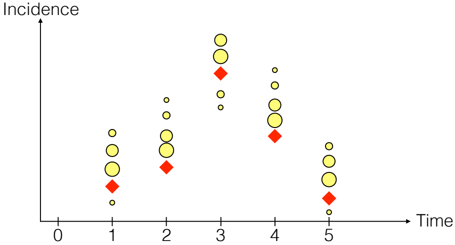
We then calculate

$$p(Y_T | X_n, \theta)$$

for each of the particles. The sum of these values is

$$\sum_{X_n} p(Y_{1:T} | X_n, \theta) p(X_n | \theta)$$

which is a sample estimate of the likelihood.



$p(y_1|\theta) \times p(y_2|y_1,\theta) \times p(y_3|y_{1:2},\theta) \times p(y_4|y_{1:3},\theta) \times p(y_5|y_{1:4},\theta)$

Log-Likelihood: $\log\{p(y_{1:T}|\theta)\} = \sum_{T} \log\{p(y_t|y_{1:t-1},\theta)\}$

pMCMC

- Once we can estimate $p(\mathrm{data}|\theta)$, we can combine this with the prior to evaluate the posterior \$p($\theta|\mathrm{data}$) for any θ.
- We can then use MCMC to sample from this -> pMCMC

Sampling from the posterior

MCMC

Sampling from the posterior

Estimating the likelihood

MCMC

SMC

PMCMC ✓

Sampling from the posterior

Estimating the likelihood

	MCMC	SMC
SMC	PMCMC ✓	SMC^2

Sampling from the posterior

Estimating the likelihood	MCMC	SMC
SMC	PMCMC	SMC^2
ABC	ABC-MCMC	ABC-SMC