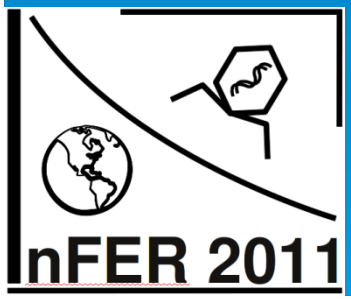
Introduction to Bayesian Analysis for Epidemic models

P.D. O'Neill
School of Mathematical Sciences
University of Nottingham



The University of Nottingham

Caveats

- This is an introduction not a masterclass!
- Aim is to present the basic ideas involved, and indicate where to find out more.

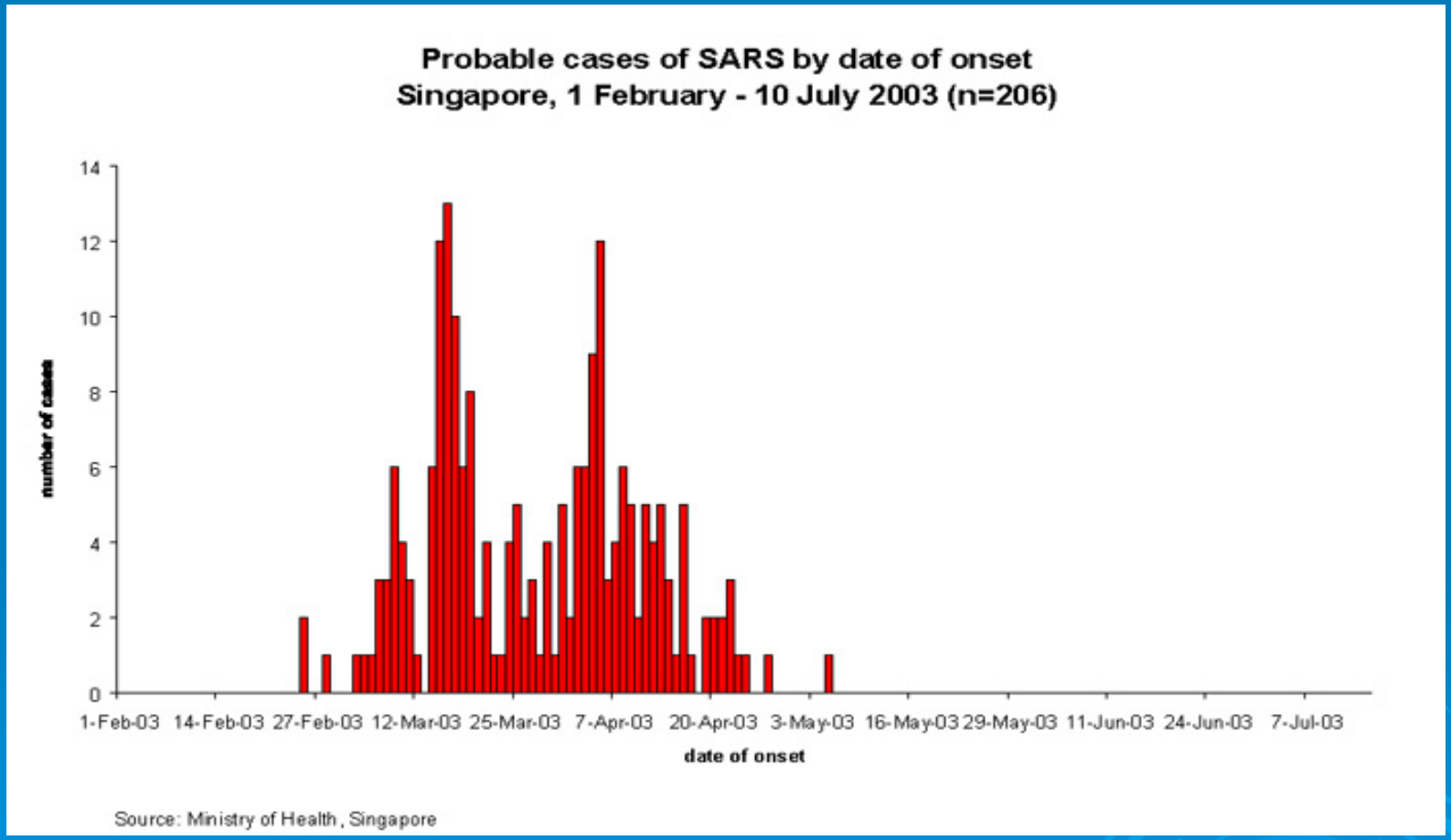
Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

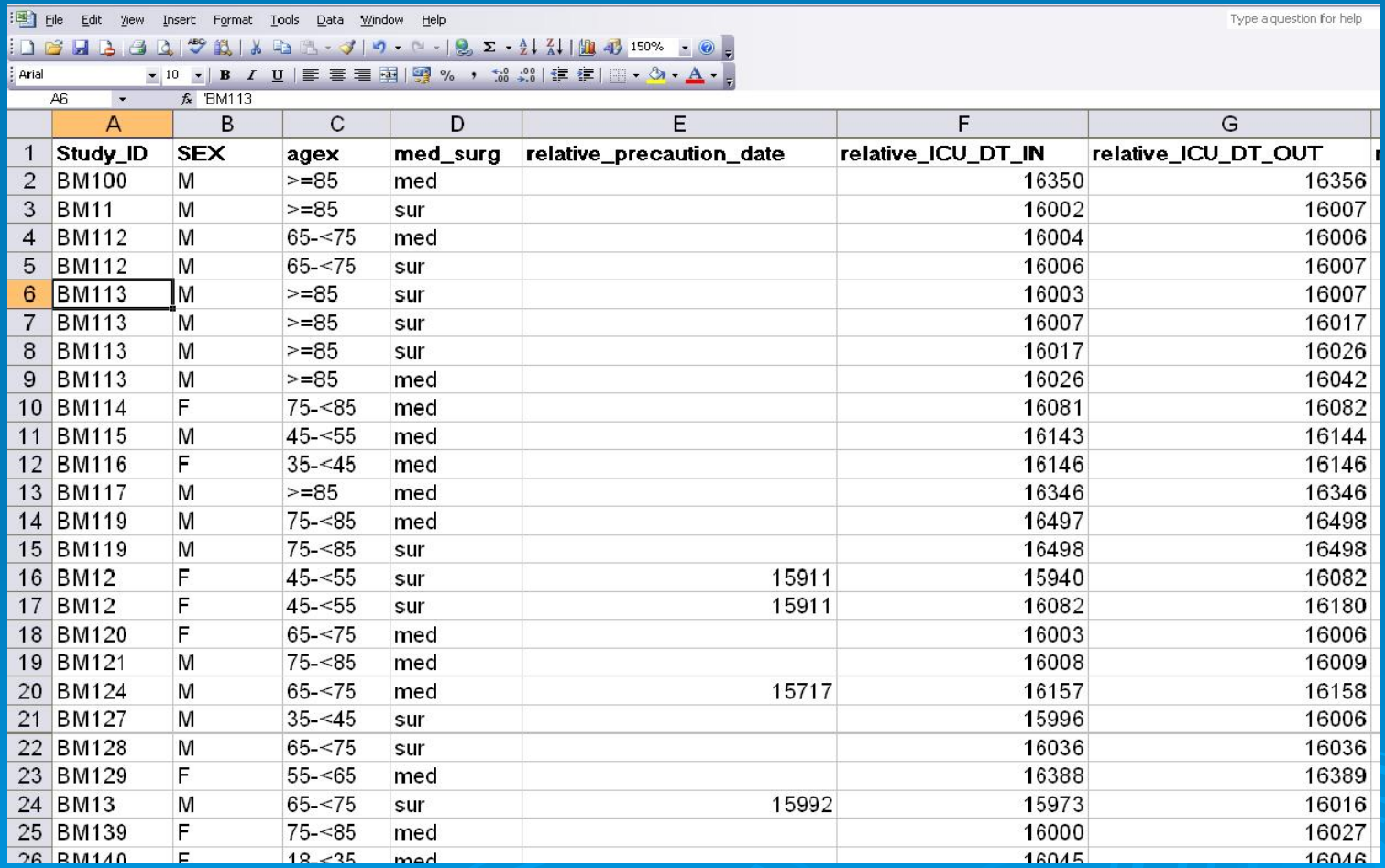
Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

1. Introduction and motivation



1. Introduction and motivation



File Edit View Insert Format Tools Data Window Help — Type a question for help

Arial 10 — 150% — A6 'BM113

	A	B	C	D	E	F	G
1	Study_ID	SEX	agex	med_surg	relative_precaution_date	relative_ICU_DT_IN	relative_ICU_DT_OUT
2	BM100	M	>=85	med		16350	16356
3	BM11	M	>=85	sur		16002	16007
4	BM112	M	65-<75	med		16004	16006
5	BM112	M	65-<75	sur		16006	16007
6	BM113	M	>=85	sur		16003	16007
7	BM113	M	>=85	sur		16007	16017
8	BM113	M	>=85	sur		16017	16026
9	BM113	M	>=85	med		16026	16042
10	BM114	F	75-<85	med		16081	16082
11	BM115	M	45-<55	med		16143	16144
12	BM116	F	35-<45	med		16146	16146
13	BM117	M	>=85	med		16346	16346
14	BM119	M	75-<85	med		16497	16498
15	BM119	M	75-<85	sur		16498	16498
16	BM12	F	45-<55	sur	15911	15940	16082
17	BM12	F	45-<55	sur	15911	16082	16180
18	BM120	F	65-<75	med		16003	16006
19	BM121	M	75-<85	med		16008	16009
20	BM124	M	65-<75	med	15717	16157	16158
21	BM127	M	35-<45	sur		15996	16006
22	BM128	M	65-<75	sur		16036	16036
23	BM129	F	55-<65	med		16388	16389
24	BM13	M	65-<75	sur	15992	15973	16016
25	BM139	F	75-<85	med		16000	16027
26	BM140	F	18-<35	med		16045	16046

1. Introduction and motivation

Number infected	Number of susceptibles per household						
	1	2	3	4	5	6	7
0	66	87	25	22	4	0	0
1	13	14	15	9	4	0	0
2		4	4	9	2	1	0
3			4	3	1	1	1
4				1	1	0	0
5					0	0	0
6						0	0
7							1

1. Introduction and motivation

Why analyse data on infectious disease?

- **Identify transmission routes**
- **Identify risk groups**
- **Prediction**
- **Assessment of control measures**
- **Estimation of disease characteristics**
- **etc...**

1. Introduction and motivation

Why use models? Why not use standard statistical methods?

- **Underlying assumptions often violated**
- **Greater realism with models**
- **Greater flexibility with models**
- **Modelling output often easier to interpret**

Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

2. The SIR epidemic model

- **Stochastic individual-based model**

Assumptions:

- **Closed population of individuals**
- **Each individual is either**
  - **Susceptible - could get disease**
  - **Infective - has it and can pass it on**
  - **Removed - no longer involved (recovered and immune, isolated, dead, etc.)**

2. The SIR epidemic model

- **Individuals can only progress in the order $S \rightarrow I \rightarrow R$**

- **Notation: At time $t \geq 0$, numbers of S, I and R are respectively denoted**

  $S(t)$, $I(t)$ and $R(t)$

- **Note that $S(t) + I(t) + R(t) = \text{constant}$**

2. The SIR epidemic model

- **Initially $S(t) = N$, $I(t) = a$, $R(t) = 0$**
- **Each infective individual remains so for a random time $T_I \sim \text{Exp}(\gamma)$**

[i.e. T_I is random with an exponential distribution with rate γ ,

$P(T \le t) = 1 - e^{-\gamma t} = 1 - \exp(-\gamma t)$]

- **T_I is called the infectious period**

2. The SIR epidemic model

- **New infections occur at rate**

$$(\beta/N)\, S(t)\, I(t)$$

- **Here "rate" means rate of a Poisson process, i.e.**

P(infection occurs in next δt time units)

$$\approx (\beta/N)\, S(t)\, I(t)\, \delta t$$

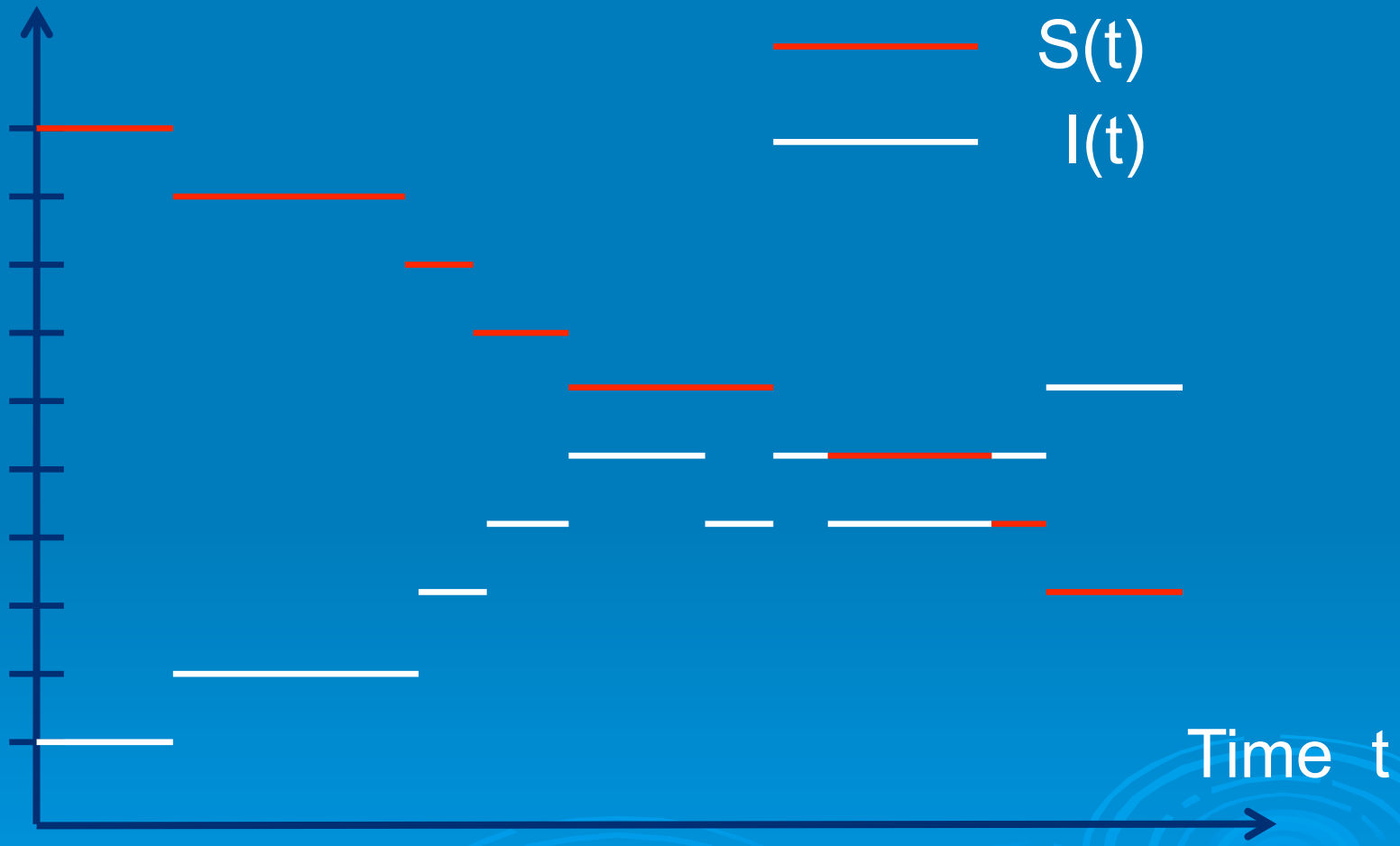
2. The SIR epidemic model

- **Each new infection means one susceptible becomes infective**
- **Epidemic ends when there are no more infectives left (since nothing else will happen)**

2. The SIR epidemic model

- **Model has two parameters:**
  - **β - the infection rate**
  - **γ - the removal rate**
- **Model has no latent period**
- **Model implicitly assumes homogeneous mixing of population**
- **All individuals are equally susceptible and infectious**

2. The SIR epidemic model



Typical realisation of stochastic SIR model

2. The SIR epidemic model

The basic model can be extended in numerous ways, e.g.

- **Different types of individual**
- **Different infectious period distributions**
- **Latent periods (SEIR, E=Exposed)**
- **SIS, SEIRS, etc**

2. The SIR epidemic model

The basic inference problem

Given data on an outbreak, we wish to estimate the model parameters

β - the infection rate

γ - the removal rate

2. The SIR epidemic model

In order to proceed, we need a likelihood,

i.e. a function f (β, γ) that tells us how probable the observed data are for any pair of values (β, γ) .

In other words,

$$f(\beta, \gamma) = P(\text{data} \mid \beta, \gamma)$$

where P(A | B) means Prob(A given B)

2. The SIR epidemic model

Maximum likelihood estimation would then proceed by finding the values of β and γ which maximised the likelihood.

In other words, the data are most likely to arise if β and γ take the ML values.

2. The SIR epidemic model

However:

- **ML estimation requires a likelihood**
- **We would like more than just point estimates. MLE theory tells us how to do this in certain cases (typically, what happens as the number of observed data becomes large) but not all**
- **Both of these can be (and are) typically a problem for the SIR model**

2. The SIR epidemic model

Likelihood = ? There are 3 basic scenarios.

1. **Complete observation through time**

   → Likelihood is known explicitly

2. **Observe removals, not infections**

   → Likelihood is intractable

3. **Just observe final number infected**

   → Likelihood can be computed, but cannot estimate both β and γ

2. The SIR epidemic model

2. ***Observe removals, not infections***

   Likelihood is intractable

This is the most common kind of data set.

ML estimation can be performed here, but becomes much harder for variants of the basic SIR model.

Instead we consider Bayesian Inference.

Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

3. Bayesian inference

Suppose we have a model with parameter vector θ and we observe data y.

Bayesian inference regards θ as an unknown random quantity in the sense that it addresses questions like

What is P(θ < 5 | y) ? What is E(θ | y)?

3. Bayesian inference

The key result is Bayes' Theorem:

$$f(\theta \mid y) = \frac{f(y \mid \theta)\, f(\theta)}{\int f(y \mid \theta)\, f(\theta)\, d\theta}$$

$f(\theta \mid y)$ = posterior density of θ given y

$f(y \mid \theta)$ = likelihood of y given θ

$f(\theta)$ = prior density of θ

3. Bayesian inference

f(θ | y) = posterior density of θ given y

This is the object of interest. If we know it then we can (in theory) answer questions about θ.

3. Bayesian inference

f(y | θ) = likelihood of y given θ

As before this tells us how likely the data are to occur given a value of θ.

f(θ) = prior density of θ

This summarises our belief about θ before seeing the data.

3. Bayesian inference

However:

$$f(\theta \mid y) = \frac{f(y \mid \theta)\, f(\theta)}{\int f(y \mid \theta)\, f(\theta)\, d\theta}$$

- **The integral can be hard to evaluate**
- **We still require a likelihood**
- **Can be hard to use $f(\theta \mid y)$ to find posterior mean, variance, etc.**

3. Bayesian inference

One way to proceed is to focus not on

→ calculating f(θ | y)

but rather on

→ producing samples from f(θ | y).

This is sampled-based inference.

3. Bayesian inference

A very useful feature of sampled-based inference is that allows us to compute summaries of $f(\theta \mid y)$ (e.g. mean, variance, correlations, ...) very easily.

For example, if we had N=1000 samples from $f(\theta \mid y)$ then we estimate $E(\theta \mid y)$ by the sample mean.

Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

4. Bayesian computation

Here we focus on

sample-based inference

and in particular

Markov chain Monte Carlo (MCMC) methods.

4. Bayesian computation

What is MCMC?

It is a method for producing (approximate) samples from a target density $g(\theta)$, where

$g(\theta)$ is known up to proportionality

In our case $g(\theta) = f(\theta \mid y) \propto f(y \mid \theta) f(\theta)$

4. Bayesian computation

What is MCMC?

Note that θ can be high-dimensional and often is in applications (as we shall see later).

4. Bayesian computation

How does MCMC work?

It works by defining a Markov Chain on the set of possible values of θ.

4. Bayesian computation

How does MCMC work?

- **A Markov chain is a random process defined in discrete time n = 0, 1, 2, 3, ...**
- **At each time point the chain randomly jumps to a new location.**
- **The choice of where to jump to only depends on the current location.**

4. Bayesian computation

How does MCMC work?

- **Suitably well-behaved Markov chains (called "ergodic") have the property that, after a long time, the distribution of their location converges to the so-called stationary or equilibrium distribution.**
- **Thus, samples from the chain after a long time are (≈) samples from the stationary distribution.**

4. Bayesian computation

A very simple example

Define a Markov chain on the states {1,2,3}

by P(1→ 1) = 0.5

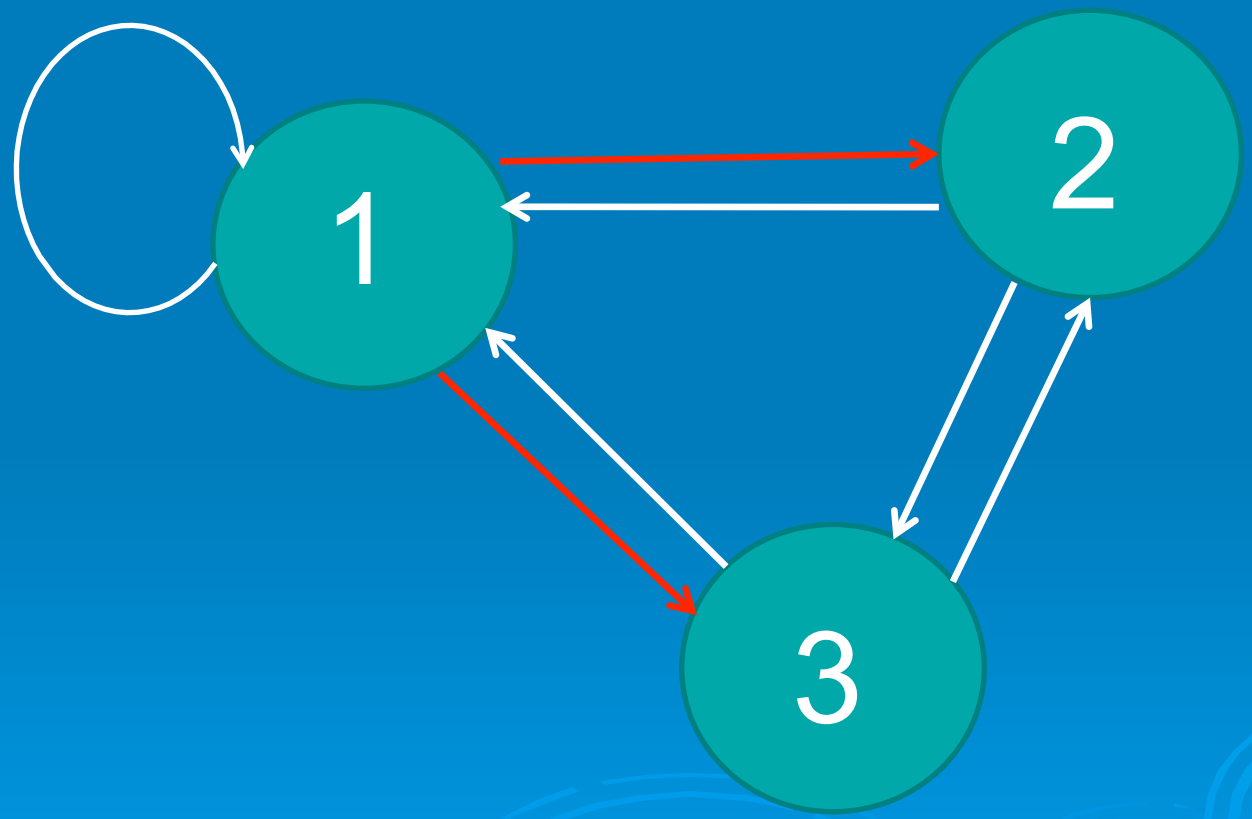
P(1→ 2) = P(1→ 3) = 0.25

P(2→ 1) = P(2→ 3) = 0.5

P(3→ 1) = P(3→ 2) = 0.5

4. Bayesian computation

A very simple example



Jump prob = 0.5

Jump prob = 0.25

4. Bayesian computation

A very simple example

Let X(n) be position of chain after n steps.

Can show that for large n

$P(X(n) = 1) \approx 0.5$

$P(X(n) = 2) = P(X(n) = 3) \approx 0.25$

So e.g. the chain spends about half its time in state 1.

4. Bayesian computation

A very simple example

So if we simulated (on a computer) the chain for e.g. 10,000 jumps and then recorded the value of X(10,001),

$$P(X(10{,}001) = 1) \approx 0.5$$

$$P(X(10{,}001) = 2) = P(X(10{,}001) = 3) \approx 0.25$$

4. Bayesian computation

Back to how MCMC works

We already have a target density $g(\theta)$ that we wish to sample from.

So the problem becomes: can we construct a Markov Chain whose stationary distribution is $g(\theta)$?

If so, then we can simulate the chain for a long time, take samples, and these will be $\approx$ samples from $g(\theta)$.

4. Bayesian computation

Metropolis-Hastings algorithm

It turns out that there are many ways to construct the required Markov chain.

The most fundamental of these is the Metropolis-Hastings algorithm.

4. Bayesian computation

Metropolis-Hastings algorithm

- Suppose chain is currently in state θ.
- Propose a new value, θ^*, from a proposal density $q(\theta^* \mid \theta)$.
- Accept proposed jump with probability

$$\frac{g(\theta^*)\, q(\theta \mid \theta^*)}{g(\theta)\, q(\theta^* \mid \theta)}$$

- Otherwise, remain at θ.

4. Bayesian computation

Metropolis-Hastings algorithm

- It can be shown that this algorithm defines a chain with stationary distribution = $g(\theta)$ (normalised)
- Choice of q is fairly arbitrary. One common choice is normal distribution centered on current value (= "Gaussian random walk")

4. Bayesian computation

Metropolis-Hastings algorithm

- If θ is multidimensional, then the components of θ can be updated individually, all at once, or in blocks
- Can sometimes update a single component "exactly" by choosing q to be equal to g - this is called Gibbs sampling (see later).

4. Bayesian computation

MCMC algorithm in practice

- **Need to run the Markov chain for a "long time" (called the burn-in period) before sampling begins**
- **Often take samples at regular intervals thereafter - called "thinning"**

4. Bayesian computation

MCMC algorithm in practice

Initialise θ, sample size = N, burn-in = B, thinning gap = G

Loop: counter = - B to N×G

Update θ

if counter > 0 and G divides counter

then output θ

End loop

4. Bayesian computation

MCMC algorithm in practice

The output of the algorithm is a sequence of values of θ which are $\approx$ samples from the required density $g(\theta)$.

4. Bayesian computation

Data Imputation

One of the most useful things about MCMC is its ability to deal with "missing data" problems.

The term "missing data" includes both literally missing observations, and also completely unobserved quantities.

4. Bayesian computation

Data Imputation

The main problem with missing data is that it can lead to an intractable likelihood.

This is especially true of mechanistic models, which may involve unobserved quantities (e.g. infection, immunity)

4. Bayesian computation

Data Imputation

To deal with missing data:

include them as if they were extra model parameters to be estimated.

4. Bayesian computation

Data Imputation

The typical situation is:

- **Observe data y**
- **Likelihood f(y | θ) is intractable**
- **Augment with missing data x, so that augmented likelihood f(x, y | θ) is tractable**
- **Run MCMC on θ and x**

Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

5. Inference for SIR model

Complete Data

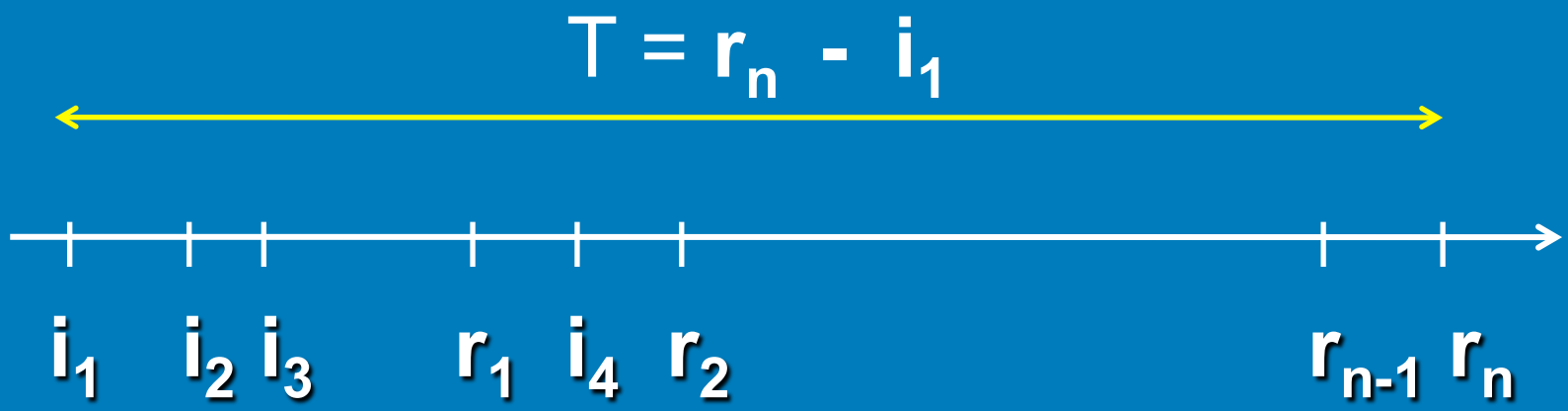
Suppose we observe the entire epidemic through time (unrealistic!)

Infection times: $i_1 \leq i_2 \leq \ldots \leq i_{n-1} \leq i_n$

Removal times: $r_1 \leq r_2 \leq \ldots \leq r_{n-1} \leq r_n$

Note i_1 is time that initial infective starts their infectious period - not modelled

5. Inference for SIR model



Infection times: $i_1 \leq i_2 \leq \ldots \leq i_{n-1} \leq i_n$

Removal times: $r_1 \leq r_2 \leq \ldots \leq r_{n-1} \leq r_n$

5. Inference for SIR model

Define $I = (i_2, \ldots, i_n)$ $R = (r_1, \ldots, r_n)$

Given β, γ, and i_1, what is the likelihood of observing I and R?

This turns out to be a standard problem from e.g. survival analysis.

5. Inference for SIR model

A heuristic argument is as follows.

- **Split time up into M very small intervals of length $\delta t = T / M$**
- **Recall that the probability of an infection occuring between t and t+ δt is $\approx (\beta/N)\, S(t)\, I(t)\, \delta t$**
- **Similarly, probability of no infection is $\approx 1 - (\beta/N)\, S(t)\, I(t)\, \delta t$**

5. Inference for SIR model

- **Multiply such probabilities together to get probability of whole infection process**
- **Let $M \to \infty$ (so $\delta t \to 0$) and convert the probability to a density (technical step)**
- **End up with:**

$$(\beta/N)^{n-1} \times S(i_2)\, I(i_2) \times S(i_3)\, I(i_3) \times \ldots \times S(i_n)\, I(i_n) \times \exp\left(- (\beta/N) \int S(t) I(t)\, dt \right)$$

5. Inference for SIR model

- **Same idea for removal process leads to**

$$\gamma^n \times I(r_1) \times I(r_2) \times \ldots \times I(r_n) \times \exp\left(-\gamma \int I(t)\, dt \right)$$

- **Overall likelihood $f(I,R \mid \beta, \gamma, i_1)$ is product of the infection and removal terms.**

5. Inference for SIR model

Bayesian inference for complete data

$$f(\beta, \gamma \mid I, R, i_1) \propto f(I,R \mid \beta, \gamma, i_1)\, f(\beta, \gamma)$$

posterior density of β and γ	likelihood	prior density of β and γ

5. Inference for SIR model

Recall that likelihood splits into infection and removal parts:

$$f(I,R \mid \beta, \gamma, i_1) =$$

$$(\beta/N)^{n-1} \times S(i_2)\, I(i_2) \times S(i_3)\, I(i_3) \times \ldots \times S(i_n)\, I(i_n)$$

$$\times \exp\left(- (\beta/N) \int S(t) I(t)\, dt \right)$$

$$\times \gamma^n \times I(r_1) \times I(r_2) \times \ldots \times I(r_n)$$

$$\times \exp\left(- \gamma \int I(t)\, dt \right)$$

5. Inference for SIR model

Thus

$$f(I,R \mid \beta, \gamma, i_1) \propto \beta^{n-1} \exp (- \beta A)$$

where

$$A = (1/N) \int S(t)I(t)\, dt$$

so if

$$f(\beta, \gamma) \propto \beta^{c-1} \exp (- \beta d)$$

then

$$f(\beta, \gamma \mid I, R, i_1) \propto \beta^{n+c-2} \exp (- \beta (A+d))$$

5. Inference for SIR model

Now if a random variable X has density

$$f(x) \propto x^{c-1} \exp(-x\,d),$$

then X has a Gamma (c,d) distribution.

Thus the posterior distribution of β is

Gamma(n+c-1, A+d)

5. Inference for SIR model

Similarly, if the prior distribution of γ is

Gamma(g,h),

then the posterior distribution of γ is

Gamma(n+g, B+h)

where $B = \int I(t)\, dt$.

Thus for complete data, posterior distributions of β, γ are standard and inference is straightforward.

5. Inference for SIR model

Incomplete data

Suppose we only observe removal times

$$R = (r_1 , \ldots , r_n)$$

Likelihood $f(R \mid \beta, \gamma)$ is practically intractable, since it involves summing over all possible infection times.

5. Inference for SIR model

We therefore use data augmentation by adding in $I = (i_2, \ldots, i_n)$ and i_1 as extra model parameters.

$$f(\beta, \gamma, i_1, I \mid R) \propto f(I, R \mid \beta, \gamma, i_1)\, f(\beta, \gamma, i_1)$$

(Note difference to complete data case !)

5. Inference for SIR model

MCMC algorithm

We need to update various parameters:

- **β, γ, i_1 can all be updated using Gibbs steps . For example**

  $$\beta \mid I, R, i_1 \sim \text{Gamma}(n+c-1, A+d)$$

  as described earlier.

5. Inference for SIR model

- **The vector of infection times**

$$I = (i_2, \ldots, i_n)$$

can be updated using a Metropolis-Hastings step.

For example, pick one infection time at random, propose a new value, and accept/reject according to M-H ratio.

5. Inference for SIR model

The whole algorithm is:

Initialise β, γ, i_1, I

Loop : update β

update γ

update i_1

update I

sample (β, γ) after burn-in

End loop

5. Inference for SIR model

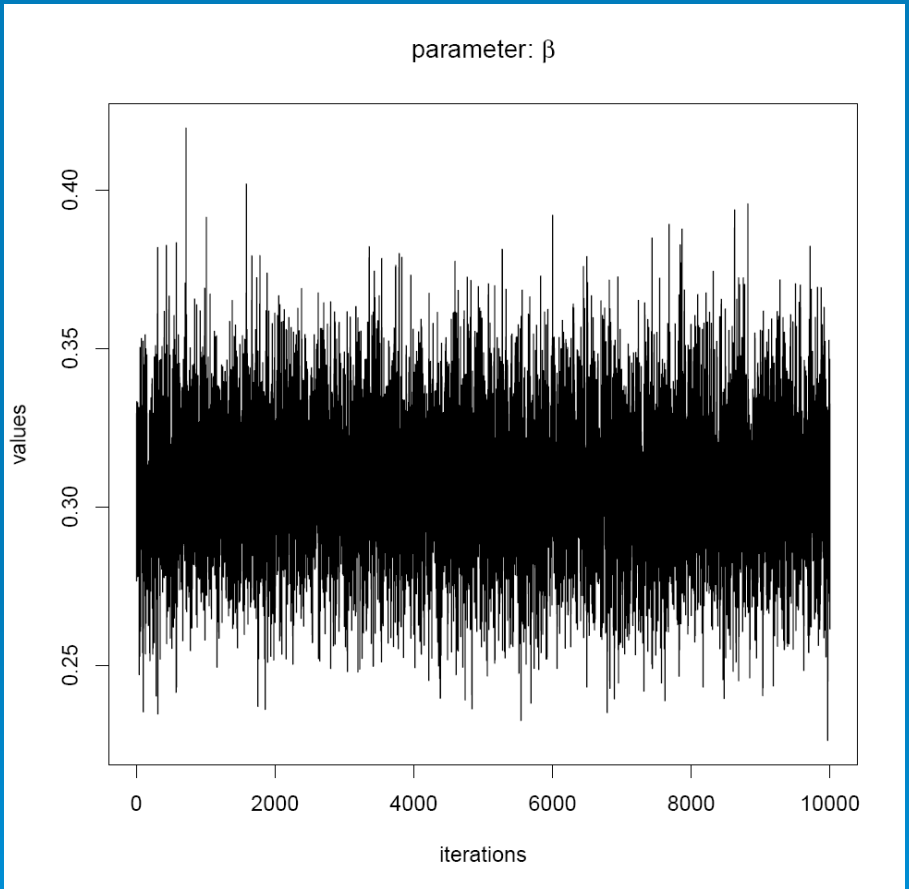
The output is a sequence of values:

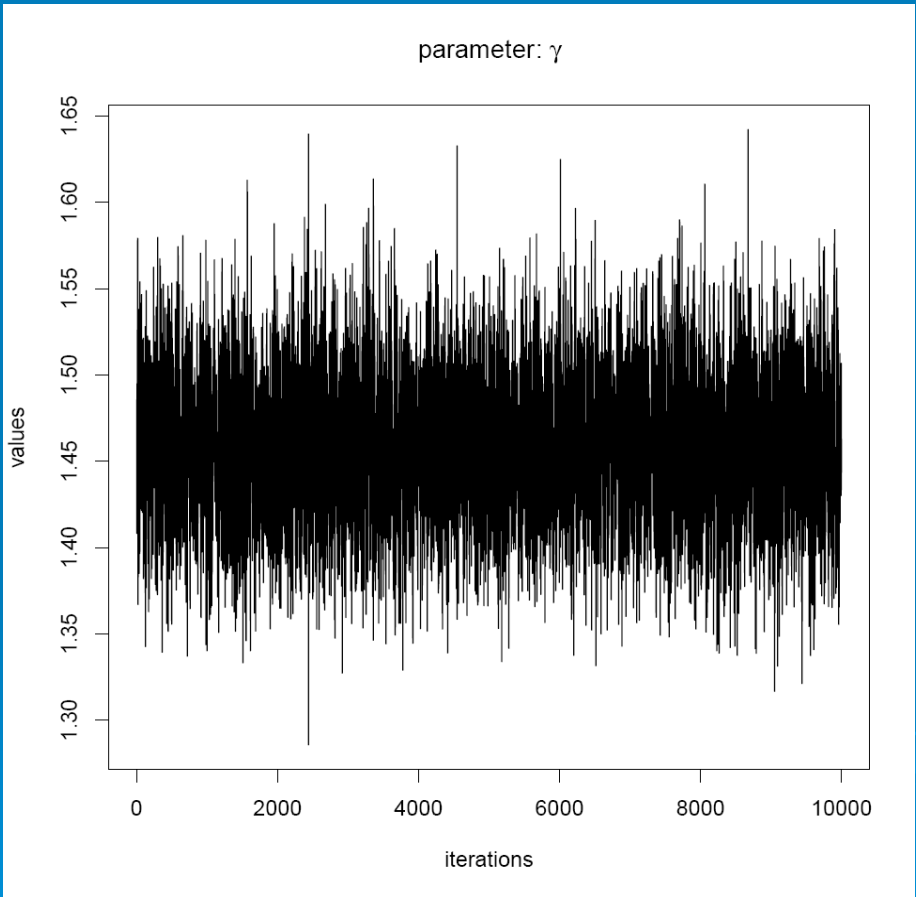
$$(\beta, \gamma)_1, (\beta, \gamma)_2, (\beta, \gamma)_3, \ldots, (\beta, \gamma)_N$$

where N = sample size, e.g. 10,000

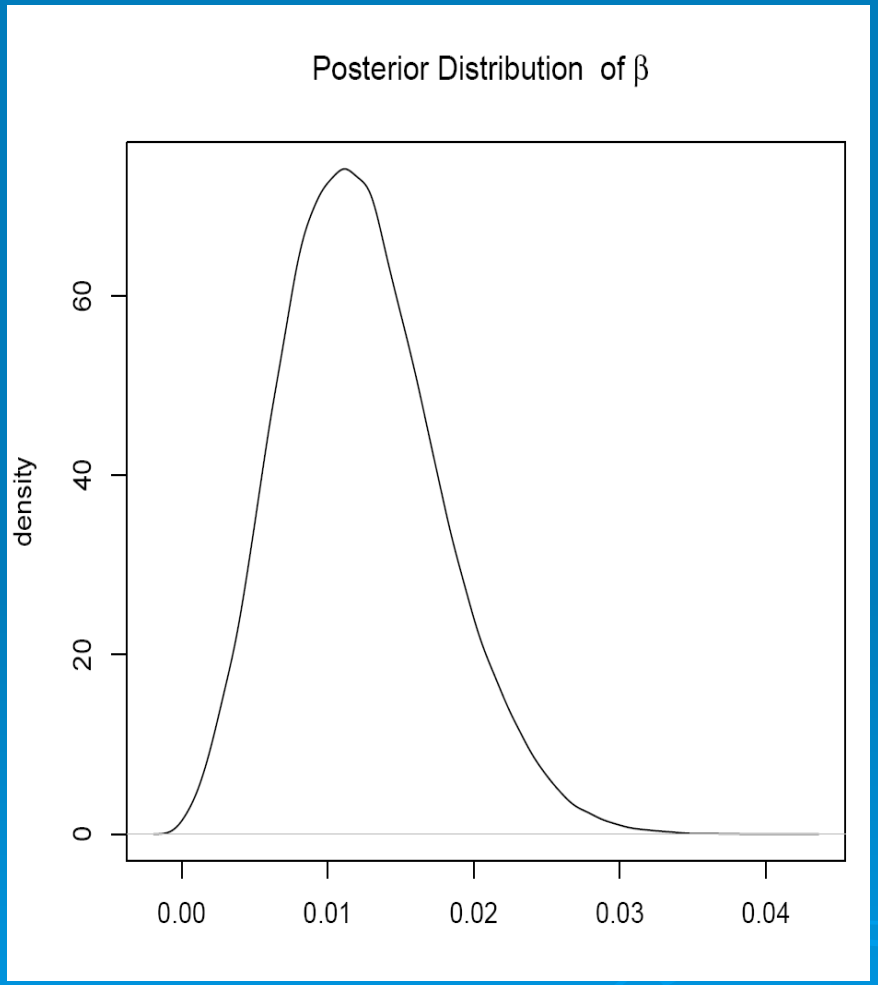
Each value is ≈ a sample from the posterior density $f(\beta, \gamma \mid R)$. The samples can be used to learn about β and γ.

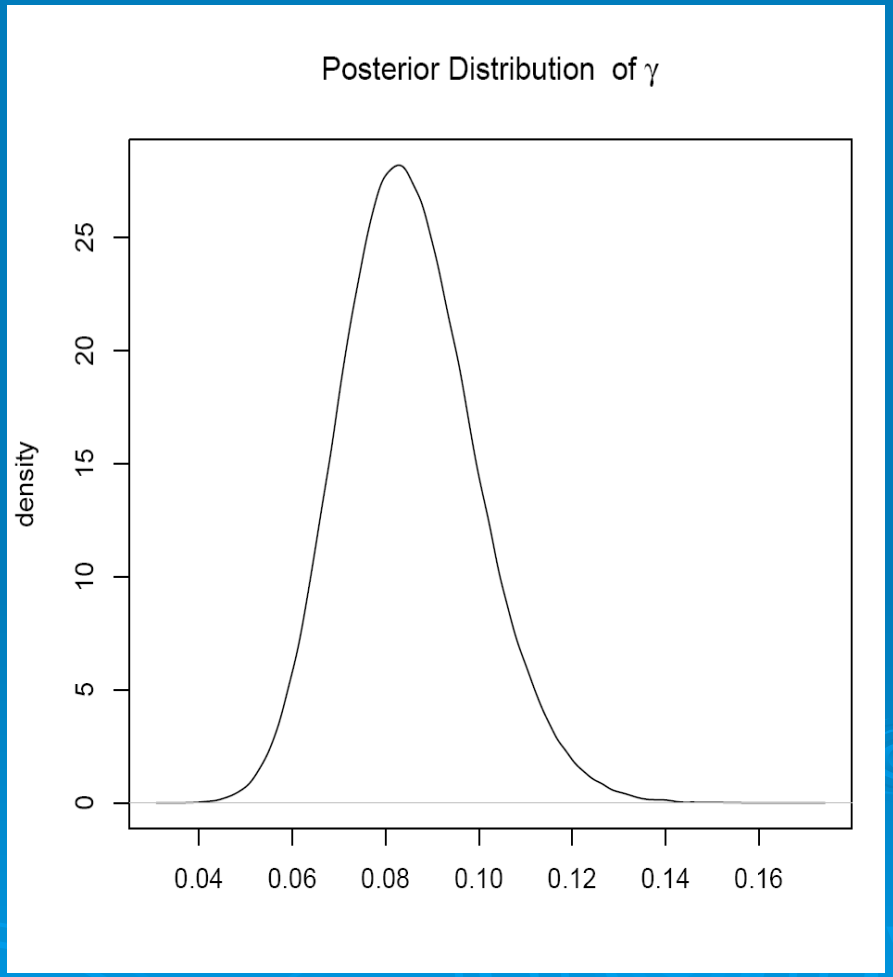
5. Inference for SIR model





5. Inference for SIR model





5. Inference for SIR model

- **It is also straightforward to obtain sampled-based estimates of other quantities, e.g.**

$$R_0 = \beta / \gamma$$

(R_0 is the basic reproduction number, which crudely is the average number of new cases generated by one index case in a large susceptible population)

5. Inference for SIR model

- In practical terms, it is often useful to test MCMC code by writing a simulation program to produce realisations of the true model
- This can then be used to test whether or not the code appears to work

Outline

1. Introduction and motivation
2. The SIR epidemic model
3. Bayesian inference
4. Bayesian computation
5. Inference for SIR model
6. Other topics

6.1 Final outcome data

Suppose we only observe the number of cases of disease in a population, or in sub-populations.

In terms of the model, this is like observing the initial conditions and the final outcome, but nothing else.

6.1 Final outcome data

1977-78 Tecumseh influenza data

Number infected	Number of susceptibles per household						
	1	2	3	4	5	6	7
0	66	87	25	22	4	0	0
1	13	14	15	9	4	0	0
2		4	4	9	2	1	0
3			4	3	1	1	1
4				1	1	0	0
5					0	0	0
6						0	0
7							1

6.1 Final outcome data

Measles outbreak in German school

	infected	total
unvaccinated	19	36
vaccinated[a]	4	782
previously infected	0	62
vaccination status not known	32	370

[a]199 were vaccinated once. Of these, two were infected.

6.1 Final outcome data

Two basic approaches to modelling household data.

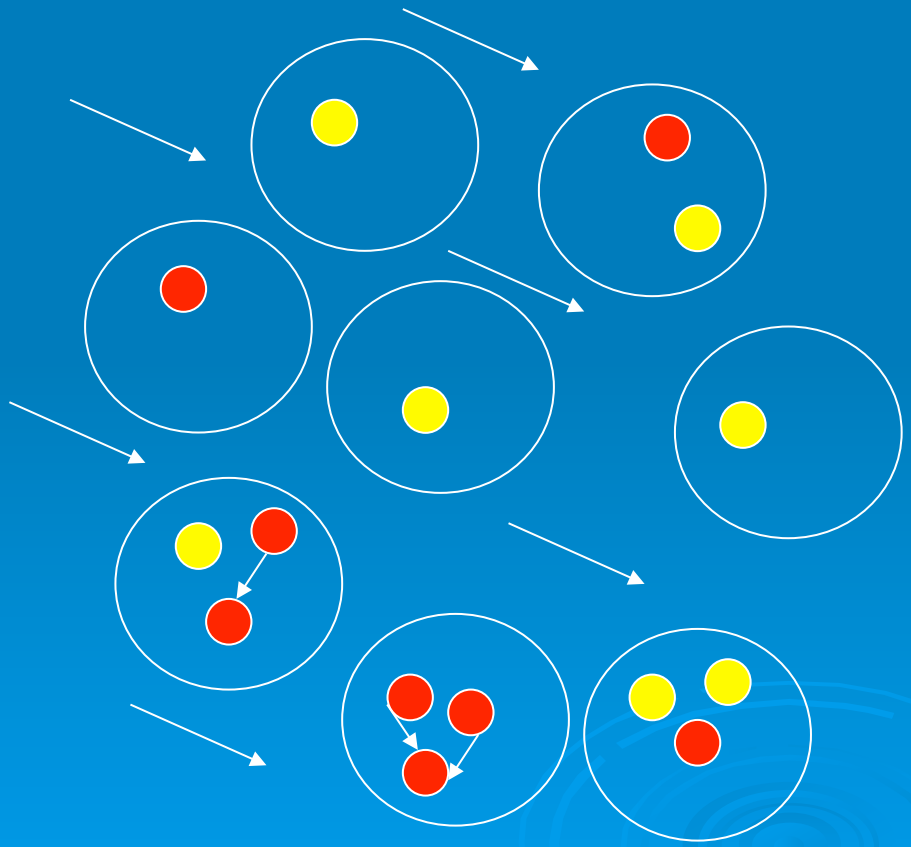
A: assume households are indepedent

B: assume interaction between households

6.1 Final outcome data

A: Independent households model

- Ever-infected
- Never-infected



SIR epidemic within each household

Each individual is subject to the same external force of infection = community infection

6.1 Final outcome data

A: assume households are indepedent

Model has two parameters, e.g.

β_H = Household infection rate

β_C = Community infection rate

Could also work with probabilities, e.g.

q_C = Community infection probability

6.1 Final outcome data

A: assume households are indepedent

If data in M households are denoted

$$y = (y_1, y_2, \ldots, y_M)$$

then likelihood is

$$f(y \mid \beta_H, \beta_C) = f(y_1 \mid \beta_H, \beta_C) \times \ldots \times f(y_M \mid \beta_H, \beta_C)$$

and each $f(y_j \mid \beta_H, \beta_C)$ **is easy to calculate**

6.1 Final outcome data

A: assume households are indepedent

Posterior density is

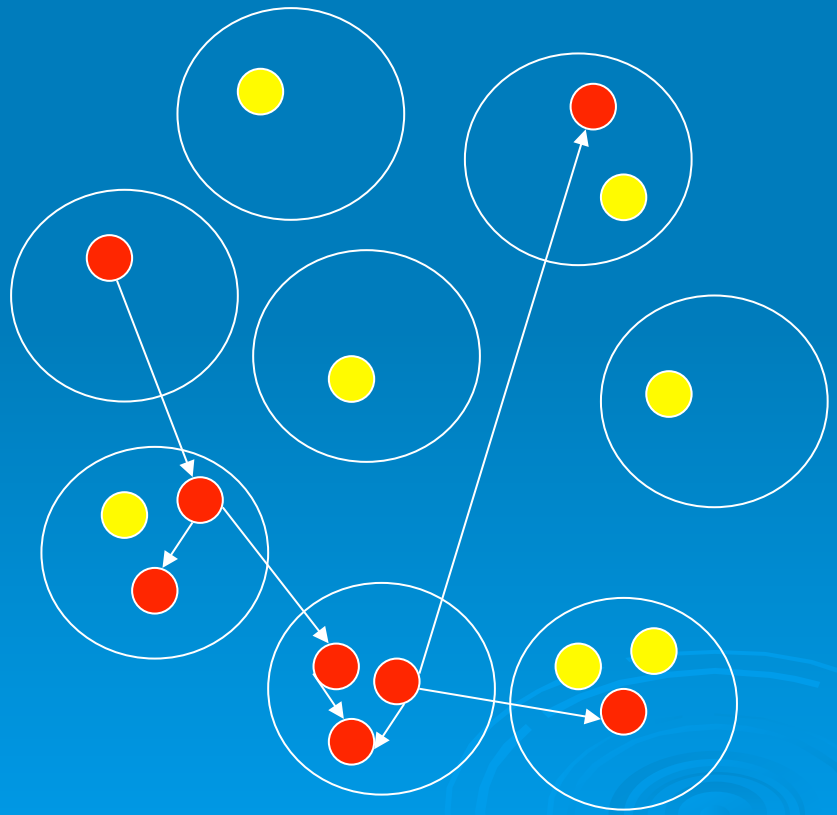
$$f(\beta_H, \beta_C \mid y) \propto f(y \mid \beta_H, \beta_C)\, f(\beta_H, \beta_C)$$

and this can be explored by MCMC or other methods (e.g. rejection sampling)

6.1 Final outcome data

B: Dependent households model (2-level mixing)

- Ever-infected
- Never-infected



SIR epidemic within each household

Individuals also interact as members of the entire community

6.1 Final outcome data

B: assume interaction between households

Model also has two parameters, e.g.

- β_H **= Household infection rate**
- β_C **= Community infection rate**

6.1 Final outcome data

B: assume interaction between households

However, now the likelihood

$$f (y \mid \beta_H, \beta_C)$$

is practically intractable since it involves summation over all possible infection paths

6.1 Final outcome data

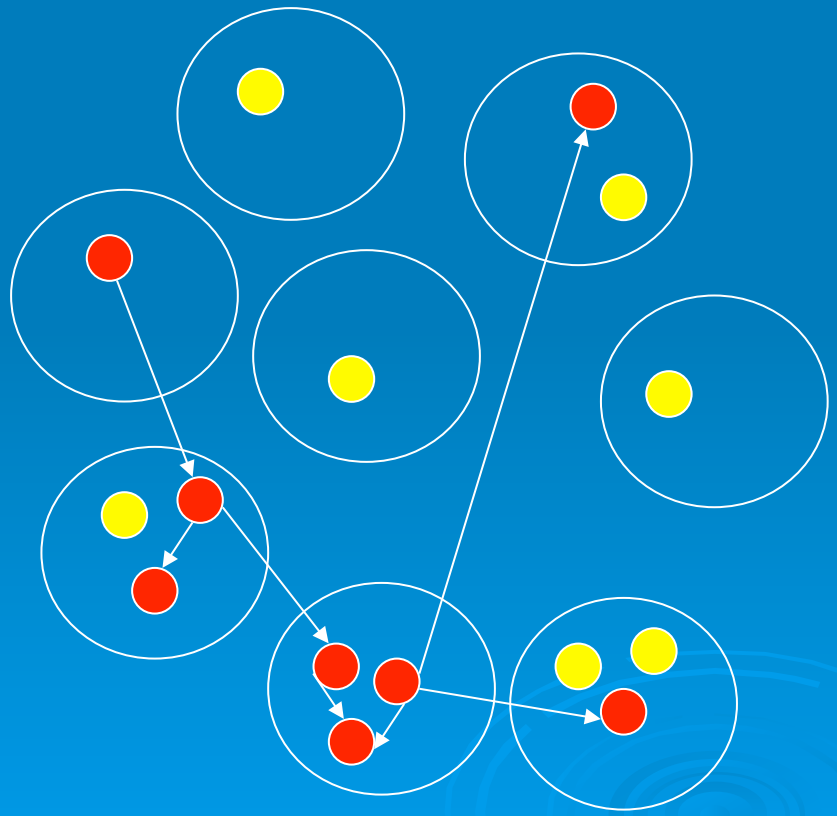
B: assume interaction between households

One solution is to use data imputation in which the missing infection process is included in the model

6.1 Final outcome data

B: Dependent households model (2-level mixing)

- Ever-infected
- Never-infected



The links: ⟶

are included as extra model parameters.

Likelihood f(y, links | β_H, β_C) is tractable

Use MCMC to update links

6.1 Final outcome data

B: assume interaction between households

The data imputation approach can be extended to take account of other missing information, e.g. vaccine status

6.2 ABC

MCMC is not the only way to be Bayesian.

- **Observe data y**
- **Propose parameter θ from prior f(θ)**
- **Simulate model given θ to obtain simulated data x**
- **If x is "close" to y (e.g. d(x,y) < ε) then accept θ**
- **θ is a sample from f (θ | d(x,y) < ε)**

6.2 ABC

- This simulation-based approach is called Approximate Bayesian Computation and is sometimes called "likelihood free inference"
- The practical challenges are choosing the distance metric d(x,y) and the tolerance ε
- Has been applied to epidemics (temporal data and final outcome data)

6.2 ABC

- ABC can also be combined with other methods: ABC-MCMC; sequential Monte Carlo ABC
- For large or complex models ABC methods can be extremely computer intensive...
- ...but it lends itself to parallel computing

6.3 Cons of MCMC

- MCMC cannot usually cope well with data-poor problems, i.e. where there is a lot of data imputation to do
- One approach to dealing with this is non-centering = reparameterising the model / algorithm to improve mixing of the Markov chain
- In some settings, lack of data makes MCMC methods infeasible

6.3 Cons of MCMC

- **MCMC can require careful algorithm design and / or tuning in order to produce results in reasonable time**
- **Assessment of chain convergence can be an issue. Although diagnostic tests exist, none are failsafe**

6.4 Model assessment

Estimating parameters is one statistical problem. But what how to assess model fit? Or to choose between models?

Various methods exist: in general the choice of method depends on the exact problem

6.4 Model assessment

Goodness of fit

Sometimes a classical statistical goodness of fit test can be used. This typically relies upon a statistic such as

“ sum of (observed - expected)^2 ”

but the underlying asymptotic theory may not always apply

6.4 Model assessment

Information Criteria

Sometimes different models can be compared using standard statistical information criteria e.g.

AIC = Akaike Information Criterion

BIC = Bayesian Information Criterion

DIC = Deviance Information Criterion

6.4 Model assessment

Bayes Factors and RJMCMC

One way to incorporate model choice into the world of MCMC is to expand the parameter space to include different models

This requires “reversible jump” MCMC (also called transdimensional MCMC)

6.4 Model assessment

Bayes Factors and RJMCMC

e.g. Suppose we havee final outcome data.

Model 1:

2-level mixing model (parameters β_H, β_C)

Model 2:

Homogeneous model (parameter β)

6.4 Model assessment

Bayes Factors and RJMCMC

- **RJMCMC works like MCMC, but now the parameter space on which the chain moves contains**

$$(M, \beta, \beta_H, \beta_C)$$

where M = 1 or M = 2 is the current model.

- **Output now includes**

P(M = 1 | data) = posterior prob that model 1 fits the data best

6.4 Model assessment

Bayes Factors and RJMCMC

- **By Bayes' Theorem,**

$$\frac{P(M = 1 \mid data)}{P(M = 2 \mid data)} = \frac{P(data \mid M=1)}{P(data \mid M=2)} \frac{P(M=1)}{P(M=1)}$$

and the ratio P(data | M=1) / P(data | M=2) is called the Bayes Factor

6.4 Model assessment

Bayes Factors and RJMCMC

- **RJMCMC is generally much harder to successfully implement than MCMC**
- **The choice of within-model prior distributions impacts the Bayes Factor (cf. Lindley's Paradox)**

6.5 Find out more

MCMC for SIR model

- **Gibson and Renshaw (1998), O'Neill and Roberts (1999)**

Stochastic epidemic models

- **Andersson and Britton (2000) (book)**

Review of methods to fit data to models

- **O'Neill (2010)**

6.5 Find out more

SISMID

Seattle,
June 2011

http://depts.washington.edu/sismid/

References

H. Andersson and T. Britton (2000) Stochastic epidemic models and their statistical analysis. Springer Lecture Notes in Statistics 151. Springer-Verlag, New York.

K. Auranen, E. Arjas, T. Leino and A. K. Takala (2000) Transmission of pneumococcal carriage in families: a latent Markov process model for binary longitudinal data. *Journal of the American Statistical Association 95*, 1044-1053.

M. Baguelin, J.R. Newton, N. Demiris, J. Daly , J.A. Mumford and J.L.N. Wood (2010) Control of equine influenza: scenario testing using a realistic metapopulation model of spread. *Journal of the Royal Society Interface* 7, 67-79.

F.G. Ball, D. Mollison and G-P. Scalia-Tomba (1997) Epidemic models with two levels of mixing. Annals of Applied Probability 7, 46-89.

N. G. Becker (1989) Analysis of Infectious Disease Data. Chapman and Hall, London.

N.G. Becker (1997) Uses of the EM algorithm in the analysis of data on HIV/AIDS and other infectious diseases. Statistical Methods in Medical Research 6, 24-37.

M. Van Boven, M. Koopmans, M. D. R. van Beest Holle, A. Meijer, D. Klinkenberg, C. A. Donnelly and H.A.P. Heesterbeek (2007) Detecting emerging transmissibility of Avian Influenza virus in human households. *PLoS Computational Biology* 3, 1394-1402.

S. Cauchemez and N.M. Ferguson (2008). Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *Journal of the Royal Society Interface* 5, 885-897.

D. Clancy and P.D. O'Neill (2007) Exact Bayesian inference and model selection for stochastic models of epidemics among a community of households. *Scandinavian Journal of Statistics* 34, 259-274.

D. Clancy and P.D. O'Neill (2008) Bayesian estimation of the basic reproduction number in stochastic epidemic models. *Bayesian Analysis*, 3(4), 737-758.

N. Demiris and P.D. O'Neill (2005) Bayesian inference for stochastic multitype epidemics in structured populations via rangom graphs. Journal of the Royal Statistical Society, Series B 67, 731-746.

References…

G. Gibson and E. Renshaw (1998). Estimating parameters in stochastic compartmental models using Markov chain methods. IMA Journal of Mathematics Applied in Medicine and Biology 15, 19-40.

M. Höhle, E. Jørgensen. and P.D. O'Neill (2005) Inference in disease transmission experiments by using stochastic epidemic models. *Applied Statistics* 54, 349-366.

M.J. Keeling, M.E.J. Woolhouse, D.J. Shaw, L. Matthews, M. Chase-Topping, D.T. Haydon, S.J. Cornell, J. Kappey, J. Wilesmith, B.T. Grenfell (2001). Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. *Science* 294, 813-817.

P.E. Lekone and B.F. Finkenstädt (2006) Statistical Inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics* 62, 1170-1177.

I.M. Longini, Jr and J.S. Koopman (1982) Household and community transmission parameters from final distributions of infections in households. *Biometrics* 38, 115-126.

T. McKinley , A.R. Cook, R. Deardon (2009) Inference in epidemic models without likelihoods. *International Journal of Biostatistics* 5(1), article 24.

P.D. O'Neill and G.O. Roberts (1999) Bayesian inference for partially observed stochastic epidemics. *Journal of the Royal Statistical Society Series A* 162, 121-129.

P.D. O'Neill, D. J. Balding, N. G. Becker, M. Eerola and D. Mollison (2000) Analyses of infectious disease data from household outbreaks by Markov Chain Monte Carlo methods. *Applied Statistics* 49, 517-542.

P.D. O'Neill (2010) Introduction and snapshot review: Relating infectious disease transmission models to data. *Statistics in Medicine* 29, 2069-2077.