Bayesian Inference for Epidemic Models

Poppy Miller and Lisa Koeppel

Lancaster University

p.miller@lancaster.ac.uk

June 9, 2017

Overview

1. Background
2. Deterministic SIR models
3. Stochastic SIR models
   - Discrete time Markov chain model
   - Continuous time Markov chain model

Why do we need to model epidemics?

- Identify the best methods of intervention via identifying covariates that increase risk of spread of the disease.
- Identifying the farms most likely to get infected next (in an ongoing epidemic)
- Estimate final size, and distribution of an ongoing epidemic to prepare resources for dealing with it
- Calculating the critical vaccination coverage
- Lots more!

Why is epidemic modelling hard?

- Strong dependencies are inherently present- the chance (hazard) that an individual gets infected depends on the status of others in their vicinity.
- The epidemic process is never fully observed
  - Models are defined in terms of who infected whom and when did this happen although genetic info is changing that.
  - We typically only observe the times that symptoms appear, therefore, true infection times are unknown.
  - The epidemic may be ongoing (therefore, not all infected individuals will have been detected yet).

The deterministic General Epidemic

[Kermack and McKendrick, 1927]

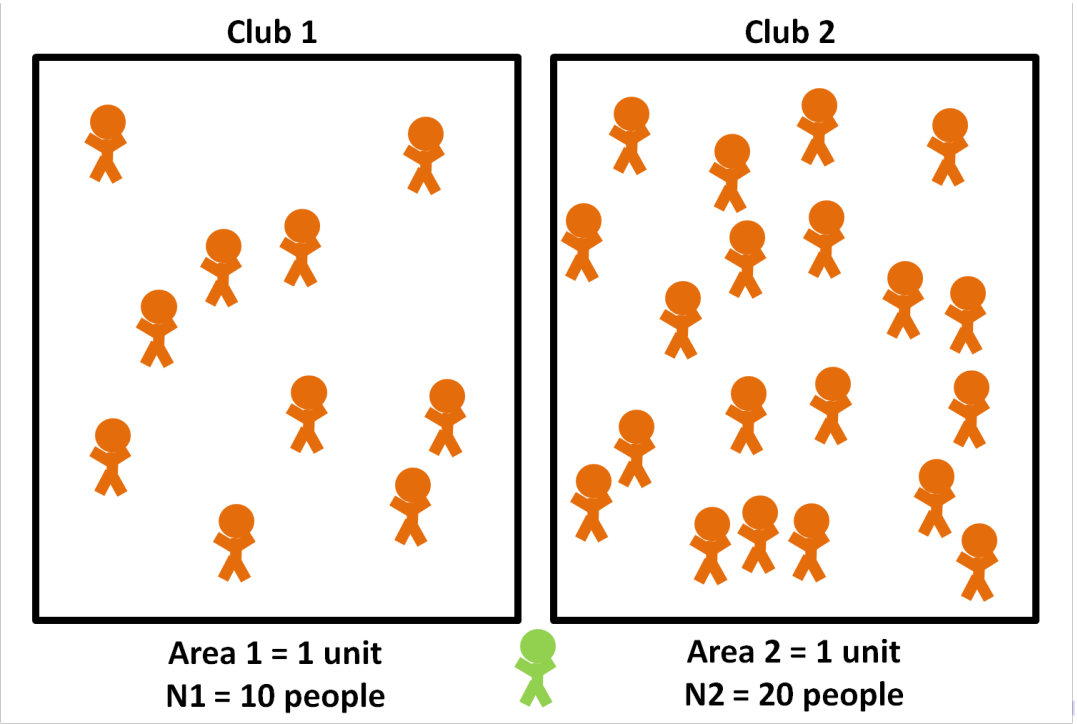
- 3 possible states: **S**usceptible → **I**nfectious → **R**emoved
- Individuals move between the states as follows:

$$\begin{aligned} \frac{dS}{dt} &= -\beta S(t) I(t) \\ \frac{dI}{dt} &= \beta S(t) I(t) - \gamma I(t) \\ \frac{dR}{dt} &= \gamma I(t) \end{aligned} \tag{1}$$

Note that $\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0$ so that the population is closed (no demography), homogeneous, and homogenously mixing. Note, we can scale the rate that susceptibles get infected to get **frequency** dependent rather than **density** dependent transmission i.e. $-\beta^* S(t) \frac{I(t)}{N}$

Frequency vs density dependent epidemics: [Unknown, 2017]

- **Density dependent:** infectious contact rate depends on the density of individuals in the study area and the prevalence of disease
  E.g. influenza transmission. $\frac{dS}{dt} = -\beta S(t) I(t)$
- **Frequency dependent:** does not depend on the density of individuals, only depends on the proportion of diseased individuals
  E.g. sexually transmitted disease transmission. $\frac{dS}{dt} = -\beta S(t) I(t) / N$



Discrete time Markov chain model: Reed-Frost model

Population: initially S_0 susceptible and I_0 infective individuals in a closed homogenously mixing population.

Model:

- Chain binomial model where events occur within a constant length time step.
- Infection duration: 1 time period
- Number of new infectives I_t follows a Binomial distribution with S_t trials and probability of success p_t.
- Probability that a given susceptible does not get infected by a given infective is q. It avoids infection from all infectives I_t with probability q^{I_t} giving $p_t = 1 - q^{I_t}$.

$$\begin{aligned} I_{t+1} &= \text{Binom}\left(S_t, p_t\right) \\ S_{t+1} &= S_t - I_{t+1} \\ R_{t+1} &= R_t + I_t \end{aligned} \tag{2}$$

Simulate from Stochastic Reed-Frost Model

```
sim_RF <- function(N, I_t, R_t, S_t, q) {
  S <- S_t
  I <- I_t
  R <- R_t
  while (I_t > 0) {
    I_new = rbinom(n = 1, size = S_t, prob = 1-q^I_t)
    S_t = S_t - I_new ## Update no in each category
    R_t = R_t + I_t   #  previous infectives recover
    I_t = I_new       #  I_new susceptibles infected
    S <- c(S, S_t)    ## Save results for
    I <- c(I, I_t)    #  this timestep
    R <- c(R, R_t)
  }
  return(data.frame(S = S, I = I, R = R))
}
```

Simulate from Stochastic Reed-Frost Model

```
data <- lapply(1:400, function(i) {
  sim_RF(N = 3, I_t = 1, R_t = 0, S_t = 2, q = 0.2)})
```

Data: final epidemic size in households of size 3 with 1 initially infective individual. Usually n_2^* and n_{21}^* are not observed but $n_2 = n_2^* + n_{21}^*$ is.

Epidemic final size and composition	**Number of households**	**Expected number of each type (n times the probability of type)**
$\boldsymbol{n_0}$ (1, 0, 0): 0	14	q^{2n_0}
$\boldsymbol{n_1}$ (1, 1, 0): 1	25	$2^{n_1}(1-q)^{n_1}q^{2n_1}$
$\boldsymbol{n_2 = n_{21}^* + n_2^*}$: 2	104 + 257	$2^{n_{21}^*}q^{n_{21}^*}(1-q)^{2n_2}$
$\boldsymbol{n_{21}^*}$ (1, 1, 1): 2	104	$2^{n_{21}^*}q^{n_{21}^*}(1-q)^{2n_{21}^*}$
$\boldsymbol{n_2^*}$ (1, 2): 2	257	$(1-q)^{2n_2^*}$

Usually we do not observe n_2^* and n_{21}^* separately, rather we observe $n_2 = n_{21}^* + n_2^*$

Bayesian Inference for Reed Frost Model

[O'Neill and Roberts, 1999]

Likelihood:

$$L\left(q; n_0, n_1, n_2, n_{21}^*\right) = 2^{n_1+n_{21}^*} q^{2n_0+2n_1+n_{21}^*} \left(1-q\right)^{n_1+2n_2}$$

Prior:

$$q \sim \text{Beta}\left(\alpha, \delta\right)$$

Marginal posteriors:

$$\pi\left(n_{21}^* | n_0, n_1, n_2, q\right) \sim \text{Binomial}\left(n_2, 2q/(2q+1)\right)$$

$$\pi\left(q | n_0, n_1, n_2, n_{21}^*\right) \sim \text{Beta}\left(2n_0 + 2n_1 + n_{21}^* + \alpha, n_1 + 2n_2 + \delta\right)$$

Use a **Gibbs sampling** scheme to sample from $\pi\left(q, n_{21}^* | n_0, n_1, n_2\right)$ by alternately drawing $q(i+1)$ and $n_{21}^*(i+1)$ from the marginal posteriors above with $n_{21}^*(i)$ and $q(i)$ respectively.

Likelihood derivation

$$L(n_0) = \prod_{i=1,\ldots,n_0} \binom{2}{0}(1-q)^0 q^2 = q^{2n_0}$$

$$L(n_1) = \prod_{i=1,\ldots,n_1} \binom{2}{1}(1-q)q \times \binom{1}{0}(1-q)^0 q = \prod_{i=1,\ldots,n_1} 2(1-q)q^2$$

$$= 2^{n_1}(1-q)^{n_1} q^{2n_1}$$

$$L(n_2^*) = \prod_{i=1,\ldots,n_2^*} \binom{2}{2}(1-q)^2 q^0 = \prod_{i=1,\ldots,n_2^*} (1-q)^2 = (1-q)^{2n_2^*}$$

$$L(n_{21}^*) = \prod_{i=1,\ldots,n_{21}^*} \binom{2}{1}(1-q)q \times \binom{1}{1}(1-q)q^0 = \prod_{i=1,\ldots,n_{21}^*} 2q(1-q)^2$$

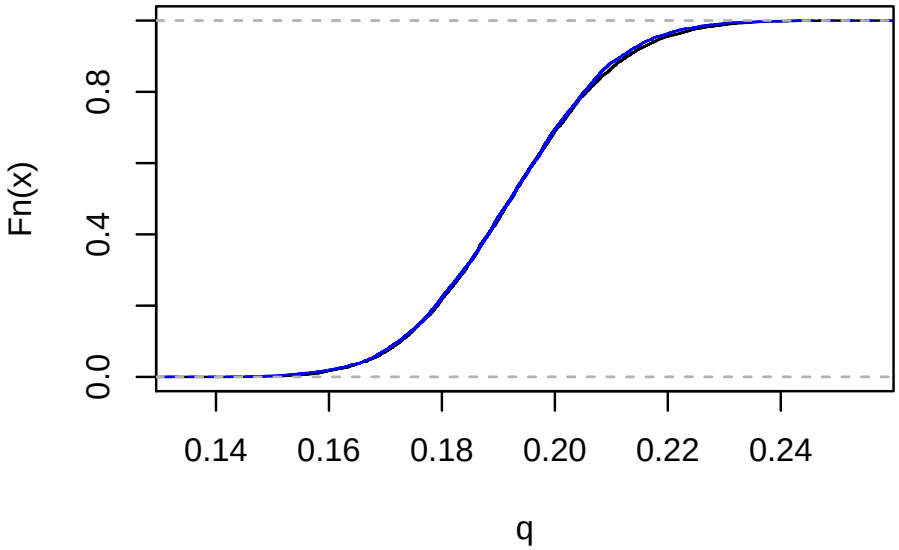
$$= 2^{n_{21}^*} q^{n_{21}^*} (1-q)^{2n_{21}^*}$$

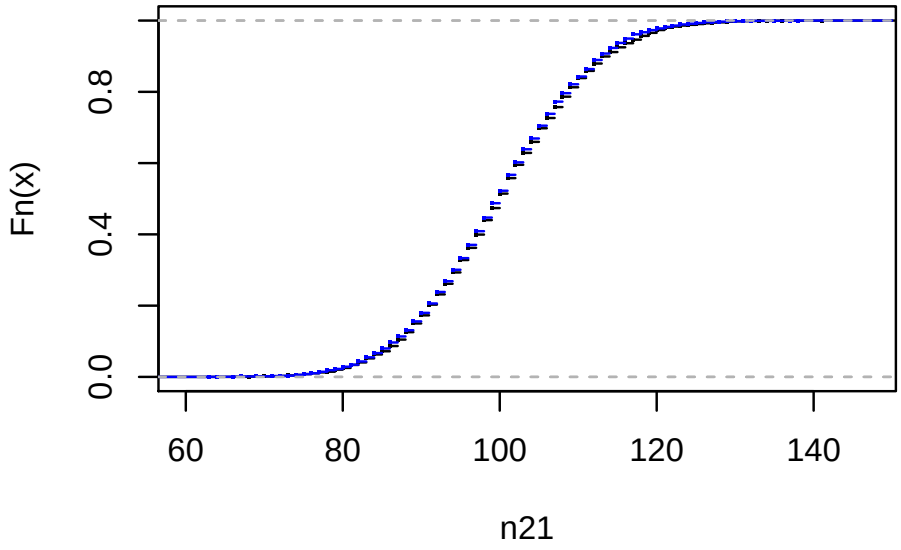
$$L(n_2) = L(n_2^*) \times L(n_{21}^*) = 2^{n_{21}} q^{n_{21}^*} (1-q)^{2(n_{21}^* + n_2^*)} = 2^{n_{21}^*} q^{n_{21}^*} (1-q)^{2n_2} \quad (3)$$

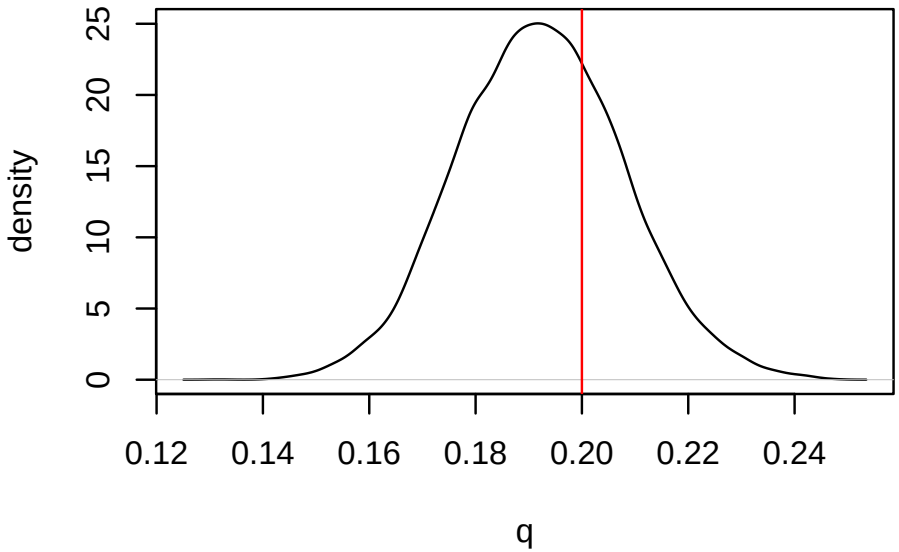
Reed Frost mcmc algorithm

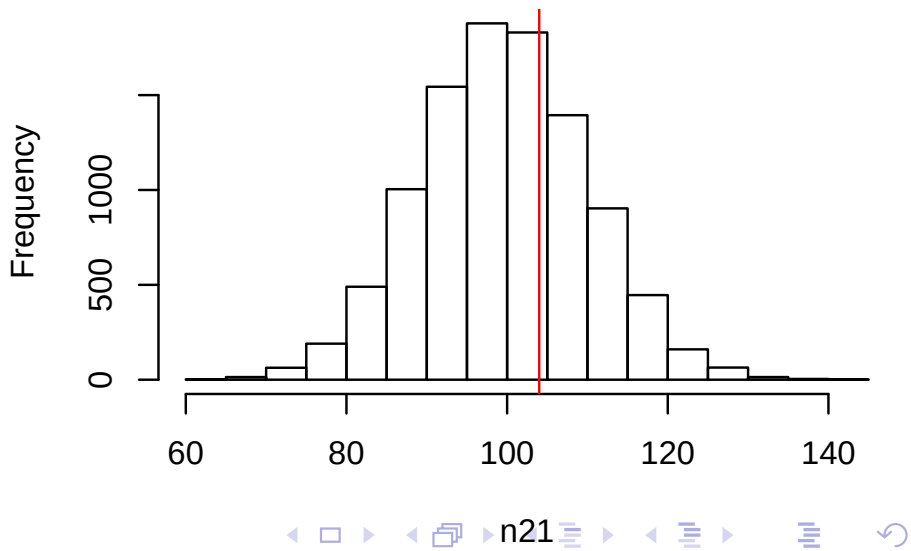
```
rf_fit <- function(n0, n1, n2, niters, inits, priors)
  q <- inits$q; n21 <- inits$n21
  out <- rbind(c(), c(q = q, n21 = n21))
  for(i in 1:niters) {
    q <- rbeta(n = 1,
               2 * (n0 + n1) + n21 + priors$alpha,
               n1 + 2 * n2 + priors$delta)
    n21 <- rbinom(n = 1, n2,
                  2 * q / (2 * q + 1))
    out <- rbind(out, c(q, n21))
  }
  return(as.data.frame(out))
}
rf_fit(n0 = 14, n1 = 25, n2 = 104+257, niters=10000,
       inits = list(q = 0.2, n21 = 104),
       priors = list(alpha = 0.01, delta = 0.01))
```

Reed Frost Results









Continuous time Markov chain model: General stochastic epidemic

Population: initially S_0 susceptible and I_0 infective individuals
Model: defined in terms of Markov transition rates:

$$\begin{aligned} P\left(\text{infection occurs in time period}[t, t+\tau_t]\right) &= \beta * S_t * I_t + o\left(\tau_t\right) \\ P\left(\text{removal occurs in time period}[t, t+\tau_t]\right) &= \gamma * I_t + o\left(\tau_t\right) \end{aligned} \tag{4}$$

Simulation: The most common method for Monte Carlo simulation of a continuous-time Markov chain is Gillespie's algorithm.
Gillespie's algorithm to simulate from SIR model

- rate of infections: $\beta I_t S_t$; rate of removals: γI_t
- simulate time till the next event from an Exponential distribution with rate the sum of the rates for all possible events.
- choose the event type with probabilities proportional to the rates for each event type

Simulate from GSE using Gillespie's Algorithm

```
sim_glsp <- function(S, I, R, N, beta, gamma) {
  transitions <- rbind(c(-1, 1, 0), c(0, -1, 1))
  tau_t <- time <- 0
  out <- rbind(c(), c(event=1, time=0, S=S, I=I, R=R))
  while (I_t > 0) {
    rates <- c(beta * I * S, gamma * I)
    tau <- rexp(1, sum(rates))
    event_type <- sample(1:2, 1, prob = rates)
    S <- S + transitions[event_type, 1]
    I <- I + transitions[event_type, 2]
    R <- R + transitions[event_type, 3]
    time <- time + tau
    out <- rbind(out, c(event_type, time, S, I, R))
  }
  return(as.data.frame(out))
}
```

Simulate from Stochastic Reed-Frost Model

```
glsp <- sim_glsp(S=29, I=1, R=0, N=30, 0.00667, 0.1)
```

Data: Removal times for each infective.

Removal times O	0, 3.6, 5.8, 7.2, 8.3, 8.7, 11.5, 11.8, 11.9, 13.2, 13.9, 13.9, 14.6, 14.7, 15.3, 17.8, 17.8, 18, 19.9, 20, 28.7, 37.1, 37.8, 45.7
Infection times Y (treated as unknown)	-10.6, -2.6, -1.4, -0.2, 2.1, 3.1, 3.2, 3.4, 4.4, 7.1, 8.3, 8.4, 9.2, 9.9, 10, 10.7, 11.2, 11.4, 11.6, 12.5, 12.5, 13.6, 18.5, 34.8

- $\mathbf{O} = (O_1, O_2, \ldots O_{n_R})$: observed removal times during $[0, T]$, $T > 0$
- $O_1 = 0$ by definition
- $\mathbf{Y} = (Y_2, Y_3, \ldots Y_{n_I})$: unobserved infection times during $(Y_1, T]$
- If the epidemic has ceased, $n_R = n_I$ else $n_R \leq n_I \leq N$
- I_t and S_t are the numbers of infected and susceptible individuals at time t

Bayesian inference for General Stochastic Epidemic (homogenously mixing

Conditional Likelihood:

$$L(\mathbf{O}, \mathbf{Y}|\beta, \gamma, Y_1) = \prod_{i=1}^{n_R} \gamma I_{O_j-} \prod_{j=2}^{n_I} \beta S_{Y_j-} I_{Y_j-} \exp\left\{-\int_{Y_1}^{T} (\beta S_t I_t + \gamma I_t)\, \mathrm{d}t\right\} \tag{5}$$

Priors:

$$\beta \sim \text{Gamma}(\lambda_\beta, \nu_\beta)\,; \gamma \sim \text{Gamma}(\lambda_\gamma, \nu_\gamma)$$

Marginal posteriors:

$$\begin{aligned} \pi(\beta|\mathbf{O}, Y_1, \mathbf{Y}, \gamma) &\sim \text{Gamma}\left(n_I - 1 + \lambda_\beta, \nu_\beta + \int_{I_1}^{T} S_t I_t \mathrm{d}t\right) \\ \pi(\gamma|\mathbf{O}, Y_1, \mathbf{Y}, \beta) &\sim \text{Gamma}\left(n_R + \lambda_\gamma, \nu_\gamma + \int_{Y_\kappa}^{T} I_t \mathrm{d}t\right) \end{aligned} \tag{6}$$

Bayesian inference for GSE (homogeneously mixing)

Updating Y_1 (first infection time):

- Prior density for Y_1 is $\theta \exp(\theta Y_1) \mathbb{1}(Y_1 < 0)$.
- The model on Y_1 is calculating the probability that the first event following Y_1 is Y_2.
- Γ is the overall rate of events multiplied by the prior rate θ.

$$\pi(Y_1|\mathbf{O}, \mathbf{Y}, \gamma, \beta) = \Gamma \exp^{-\Gamma(Y_2 - Y_1)}$$

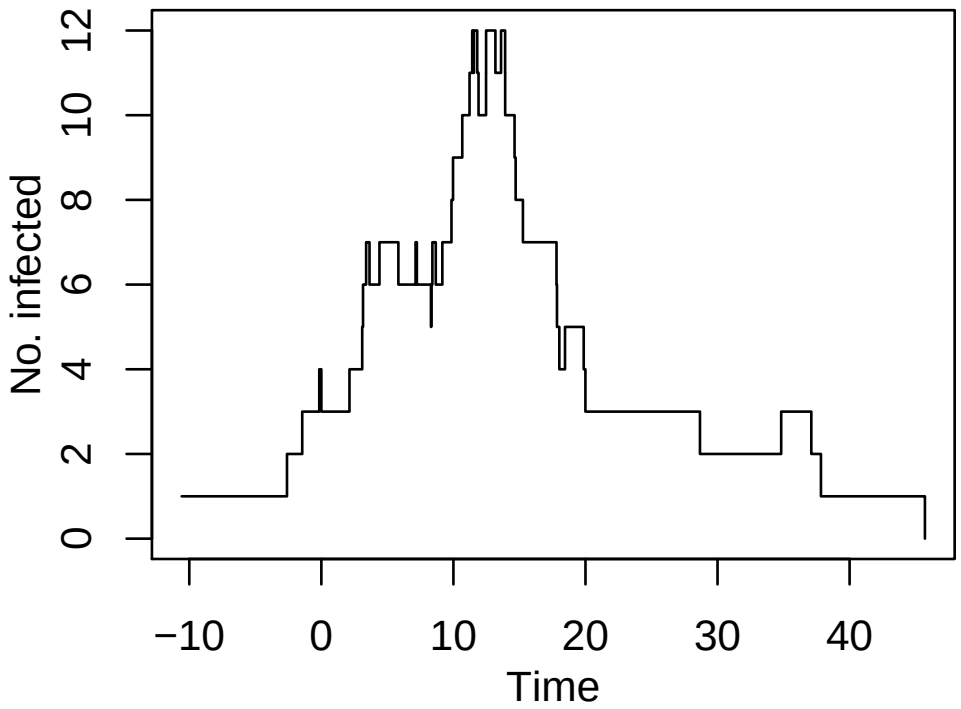
where $\Gamma = \theta + \gamma + \beta N$.

- This is equivalent to $Y_2 - Y_1 \sim$ Exponential(Γ).
- Therefore, sample $\xi \sim \text{Exp}(\Gamma)$, and set $Y_1 = Y_2 - \xi$.

Bayesian inference for GSE (homogeneously mixing)

Calculating integrals:

$$\int_{I_1}^{T} I\,(\mathrm{t})\,\mathrm{d}t = \sum_{i=2}^{T} (t_i - t_{i-1})\, I_{t_i}^{-}$$



Updating **Y** using reversible jump

Updating Y: Sampling from $\pi(\mathbf{Y}|\mathbf{O}, Y_1, \beta, \gamma)$ is done via a reversible jump hastings algorithm with 3 steps (move infection time, add new infection time, remove infection time).

- **Move** an infection time
  One of the existing infection times (s) is chosen uniformly at random, and replacement time (t) is obtained by sampling Uniformly on $[Y_1, T]$. The new time is accepted with probability: $\frac{f(Y-\{s\}+\{t\})}{f(Y)} \wedge 1$
- **Add** an infection time
  A new time (t) is Uniformly sampled from (Y_1, T) and added to **Y**. The new infection is accepted with probability: $\frac{f(Y+\{t\})(T-Y_1)}{f(Y)(n+1)} \wedge 1$
- **Remove** an infection time
  An existing infection time (s) is randomly (uniformly) selected to be removed from **Y**. The new infection is accepted with probability: $\frac{f(Y-\{s\})n}{f(Y)(T-Y_1)} \wedge 1$

Problems with updating I using reversible jump

Given N is large, the epidemic is not concluded and the number of removed individuals is low:

- The number of possible infectives is huge which results in very slow exploration of the sample space using the reversible jump algorithm.
- The algorithm is already very slow when modelling heterogeneously mixing population.

Bayesian inference for GSE (homogenously mixing)

[Britton and O'Neill, 2002]

Different (but equivalent) setup to the previous epidemic. This one is more easily extended to heterogenously mixing.

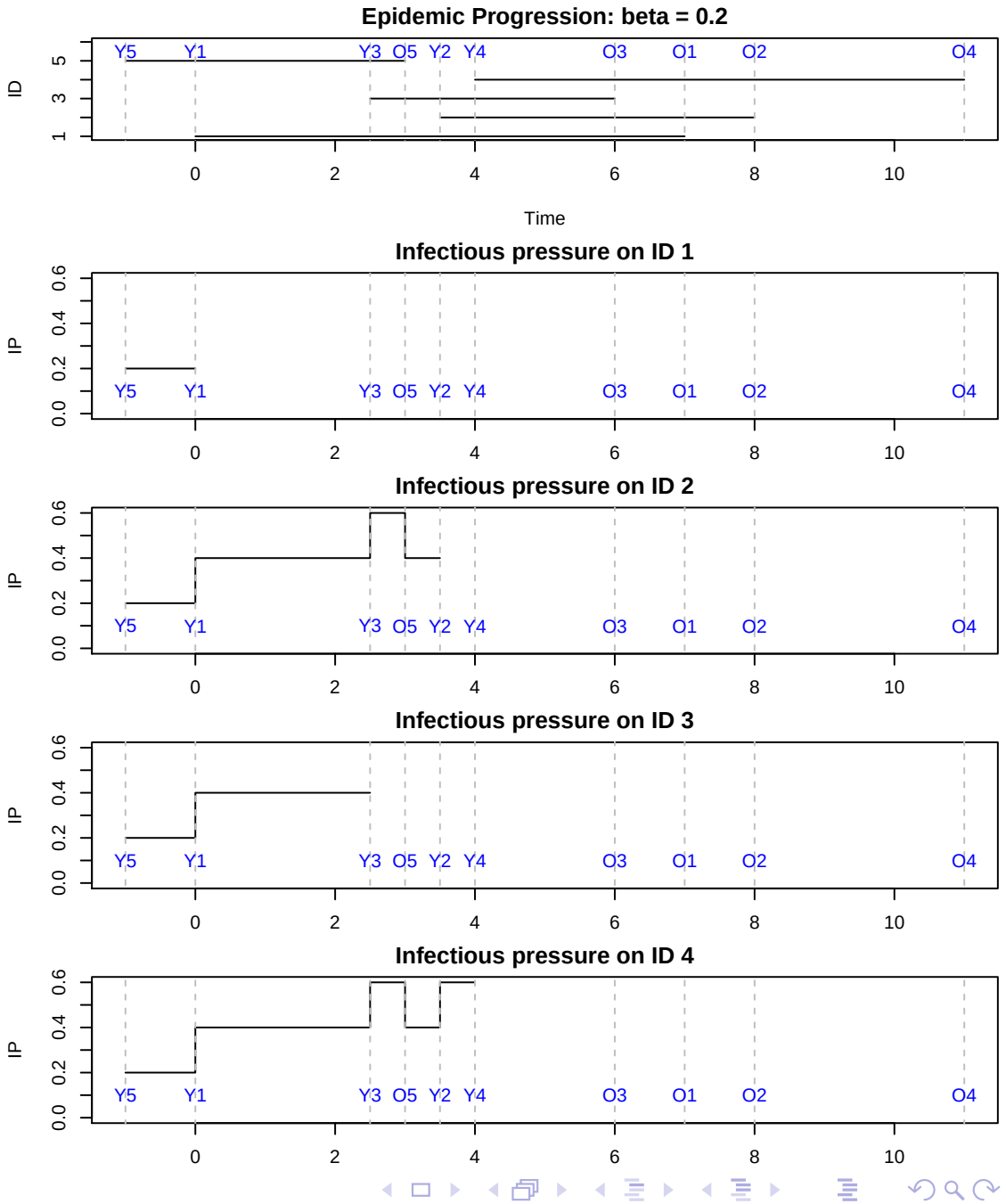
- Label infected individuals: $1, ..., n_I$ (initial infective is labeled κ)
- Label uninfected individuals: $n_I + 1, ..., N$.
- Each individual i gets an infection time Y_i and a removal time O_i (where $Y_i = O_i = \infty$ for $i > n_I$).
- A susceptible individual j (when it becomes infected) gets infectious pressure β from an infected individual i iff $Y_i < Y_j < O_i$.
- Total infectious pressure on i when it gets infected at time t_i is $P_i = \sum_{j \in I_{t_i}} \beta$

Total infectious pressure on all individuals in epidemic:

$$G = \int_{Y_\kappa}^{T} \beta I_t S_t \mathrm{d}t$$
$$= \int_{Y_\kappa}^{T} \sum_{i \in I_t} \sum_{j \in S_t} \beta \mathrm{d}t$$
$$= \sum_{i=1}^{n_I} \sum_{j=1}^{N} \beta \left((O_j \wedge Y_i) - (Y_j \wedge Y_i) \right)$$

Example for ID 4:

$$\begin{aligned} G_{i=4} &= \beta \left((O_1 \wedge Y_4) - (Y_1 \wedge Y_4) \right) \\ &+ \beta \left((O_2 \wedge Y_4) - (Y_2 \wedge Y_4) \right) \\ &+ \beta \left((O_3 \wedge Y_4) - (Y_3 \wedge Y_4) \right) \\ &+ \beta \left((O_5 \wedge Y_4) - (Y_5 \wedge Y_4) \right) \\ &= \beta \left(Y_4 - Y_1 \right) + \beta \left(Y_4 - Y_2 \right) \\ &+ \beta \left(Y_4 - Y_3 \right) + \beta \left(O_5 - Y_5 \right) \end{aligned}$$



Conditional Likelihood:

Infectious part: Total person-to-person infectious pressure during the course of the epidemic:

$$G = \sum_{i=1}^{n_I} \sum_{j=1}^{N} \beta \left((O_j \wedge Y_i) - (Y_j \wedge Y_i) \right)$$

giving $L_1 = \prod_{i=1, i \neq \kappa}^{n_I} \left(\sum_{j \in I_{t_i}} \beta \right) \times \exp\{-G\}$

Removal part: The infectious period for individual i is $O_i - Y_i$ giving $L_2 = \prod_{i=1}^{n_R} \gamma \exp\{-\gamma (O_i - Y_i)\}$

Note: The exponential term in L_1 states the probability that no event occurs between your last observation of an event and the end of the observation period.

Let n_I and n_R be the number of infected and removed individuals resp.

Conditional Likelihood:

$$L(\mathbf{Y}, \mathbf{O}|\beta, \gamma) = \prod_{i=1, i\neq\kappa}^{n_I} \left(\sum_{j \in I_{t_i}} \beta \right) \times \exp\{-G\} \times \gamma^{n_R} \prod_{i=1}^{n_R} \exp\{-\gamma (O_i - Y_i)\} \tag{7}$$

Priors: $\beta \sim \text{Gamma}(\lambda_\beta, \nu_\beta)$; $\gamma \sim \text{Gamma}(\lambda_\gamma, \nu_\gamma)$

Marginal posteriors:

$$\pi(\beta|\mathbf{O}, Y_\kappa, \mathbf{Y}, \gamma) \sim \Gamma\left(n_I - 1 + \lambda_\beta, \nu_\beta + \sum_{i=1}^{n_I} \sum_{j=1}^{N} ((O_j \wedge Y_i) - (Y_j \wedge Y_i)) \right)$$

$$\pi(\gamma|\mathbf{O}, Y_\kappa, \mathbf{Y}, \beta) \sim \Gamma\left(n_R + \lambda_\gamma, \nu_\gamma + \sum_{i=1}^{n_R} (O_i - Y_i) \right) \tag{8}$$

Bayesian inference for GSE (extension) [Jewell et al., 2009]

Infectious periods: The time from infection to remove of individual i is a random variable $D_i = O_i - Y_i$.
Let $f_D(x)\,(x \geq 0)$ denote the probability density function of D and

$$F_D(x) = \int_x^\infty f_D(y)\,\mathrm{d}y$$

The homogenously mixing GSE used exponentially distributed infectious periods, but any non-negative probability distribution can be used.

Bayesian inference for GSE (heterogenously mixing)

β_{ij} **parameters:** Any heterogeneity in infectiousness or susceptibility of individuals can be incorporated here. For example, spatial kernel for distance between individuals, individual covariates, contact networks between individuals etc. Easiest when $\beta_{ij} = \beta_0 h_{ij}$ where h_{ij} is fully known.

Conditional Likelihood:

$$L(\mathbf{Y}, \mathbf{O}|\beta, \gamma) = \prod_{j=1, j\neq\kappa}^{n_I} \left(\sum_{i \in I_{t_j-}} \beta_{ij} \right) \times \exp\left\{ -\sum_{i=1}^{n_I} \sum_{j=1}^{N} \left((O_i \wedge Y_j) - (Y_j \wedge Y_i) \right) \beta_{ij} \right\} \times \prod_{i=1}^{n_I} f_D(O_i - Y_i) \tag{9}$$

MLE inference for GSE (homogenously mixing, epidemic completed) [Kypraios, 2007]

By differentiating the likelihood (for either setup), we get

$$\hat{\beta} = \frac{n_I - 1}{\int_{I_\kappa}^{T} S_t I_t \mathrm{d}t} \text{ or } \hat{\beta} = \frac{n_I - 1}{\sum_{i=1}^{n_I} \sum_{j=1}^{n_R} (R_i \wedge I_j - I_j \wedge I_i)} \tag{10}$$

$$\hat{\gamma} = \frac{n_R}{\int_{I_\kappa}^{T} I_t \mathrm{d}t} \text{ or } \hat{\gamma} = \frac{n_R}{\sum_{i=1}^{n_R} (R_i - I_i)} \tag{11}$$

By differentiating the first derivative of the likelihood with respect to β and γ, the corresponding standard errors turn out to be:

$$\sigma_{\hat{\beta}} = \frac{\beta}{\sqrt{n_I - 1}} \tag{12}$$

$$\sigma_{\hat{\gamma}} = \frac{\gamma}{\sqrt{n_R}} \tag{13}$$

Requires known infection times!


Britton, T. and O'Neill, P. (2002).
Bayesian inference for stochastic epidemics in populations with random social structure.
Scandinavian Journal of Statistics: Theory and Applications.


Jewell, C., Kypraios, T., Neal, P., and Roberts, G. (2009).
Bayesian analysis for emerging infectious diseases.
Bayesian Analysis.


Kermack, W. and McKendrick, A. (1927).
A contribution to the mathematical theory of epidemics.
Proceedings of the Royals Society A.


Kypraios, T. (2007).
Efficient bayesian inference for partially observed stochastic epidemics and a new class of semi-parametric time series models.
Master's thesis, Lancaster UNiversity.


O'Neill, P. and Roberts, G. (1999).
Bayesian inference for partially observed stochastic epidemics.
Journal of the Royal Statistical Society: Series A (Statistics in Society), 162(1):121–129.


Unknown (2017).
Density-dependent vs. frequency-dependent disease transmission.