Non-parametric Bayesian Inference for Epidemic Models

Rowland Seymour

University of Nottingham

Meeting on Stochastic Epidemic Models, July 2017

Table of Contents

Table of Contents

SIR Models

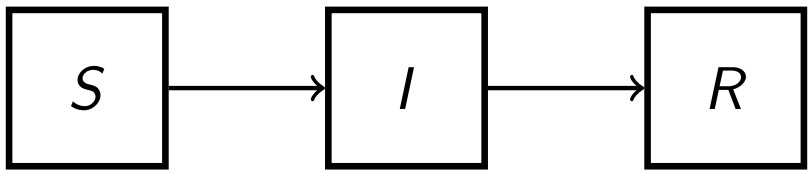


Figure: A typical SIR diagram.

Given infection and removal times, can we work out the rate that individuals move between classes?

In the standard epidemic model, we assume infections occur according to a Poisson process with rate $\beta_0 S_t I_t$ and the removals occur according to a Poisson process with rate γI_t.

In epidemic inference, we try to estimate β_0 and γ.

The Infection Rate

In the standard epidemic model, the infection rate is $\beta_0 S_t I_t$.

Problems:

- Is this the functional form of the infection rate?
- Do other covariates, such as time or distance, affect the infection rate?

Solutions:

- Use non-parametric inference.
- Inference for a heterogeneously mixing model.

Table of Contents

Non-Parametric Models

Using a non-parametric inference has the following advantages:

- Let the data speak for itself
- Remove bias from model choice
- Flexible choice of model

Assumptions and Ideas

Infections:
Instead of assuming the infection rate is $\beta_0 S_t I_t$, we assume the infection rate can be modelled by an inhomogeneous Poisson process with rate β, where

$$\beta = f(t) \quad \text{or} \quad \beta = f(x, y) \quad \text{or} \quad \beta = f(S_t, I_t).$$

Recoveries:
The infected individuals remain so for a period, which has some given distribution, for example:

$$\text{Exp}(\gamma) \quad \text{or} \quad \text{Gamma}(\lambda, \gamma).$$

Bayesian Inference for Non-Parametric Models

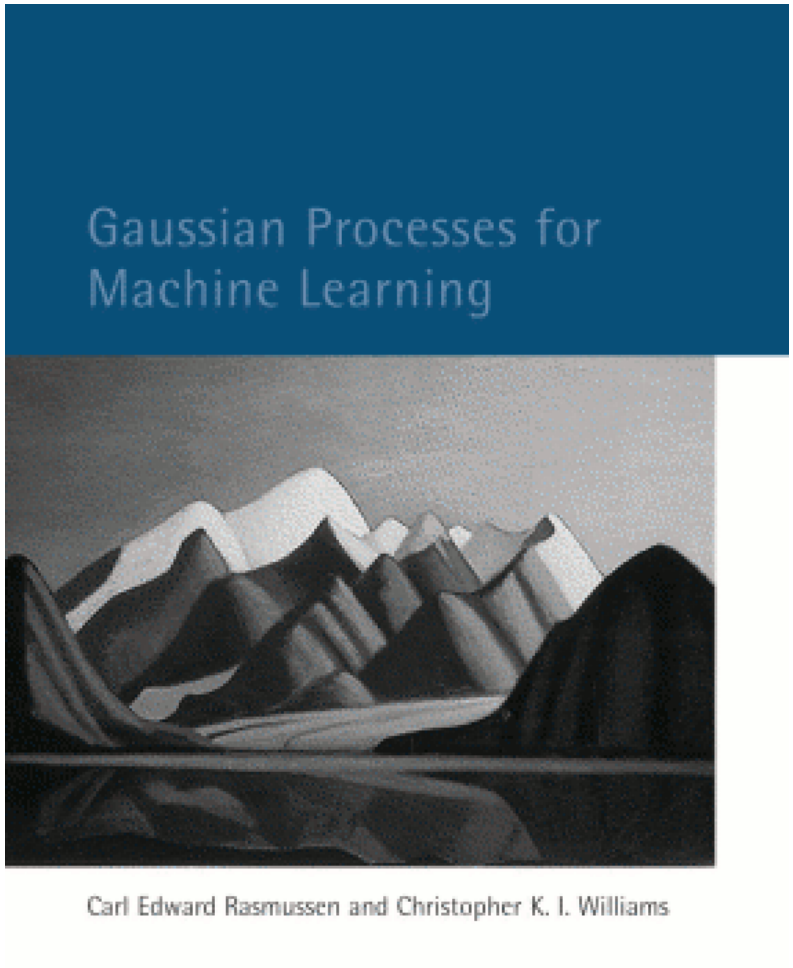
Problems:
We have an infinite set of functions for β. How do we place a prior distribution on β and infer the function in a reasonable amount of time?

Solutions:
We use Gaussian processes to estimate the infection kernel.

Table of Contents

What are Gaussian Processes?

- Gaussian processes (GPs) are a popular machine learning tool for learning functions.
- The Multivariate Gaussian distribution is over vectors, and GPs are over functions.
- One of the main uses in non-parametric regression.



Defintion

Definition

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.

To specify a GP, we need to define the mean $m(\mathbf{x})$ and covariance function $k(\mathbf{x}, \mathbf{x}')$. We write it as

$$f \sim \mathcal{GP}\big(m(\mathbf{x}),\ k(\mathbf{x}, \mathbf{x}')\big)$$

We can input our assumptions of the function through the covariance function.

Covariance Functions

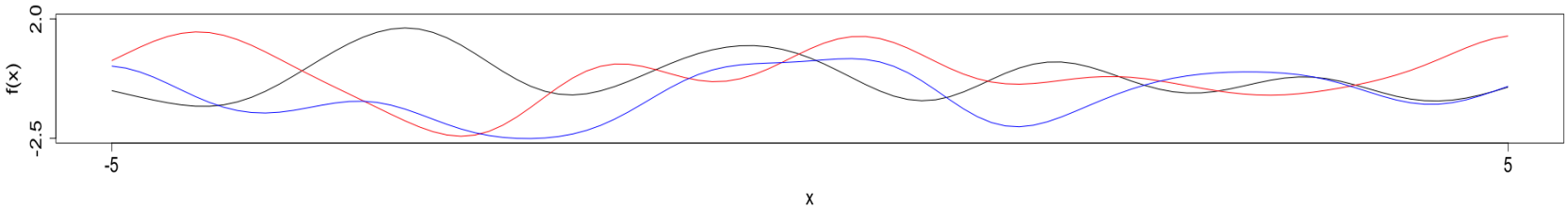


Figure: Three draws from the square exponential covariance function.

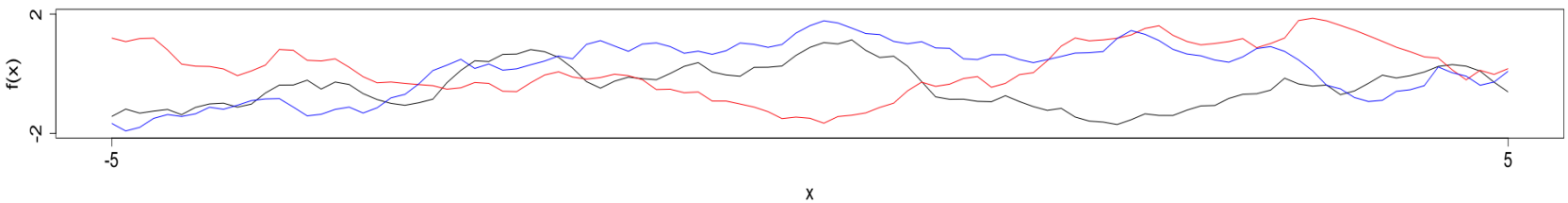


Figure: Three draws from the Mátern covariance function.

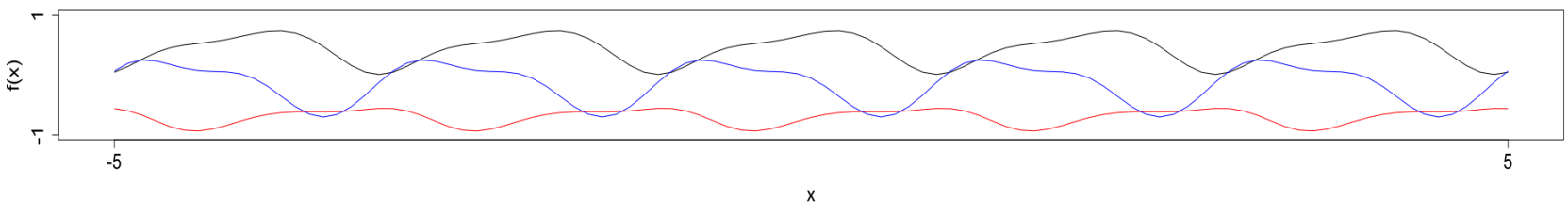
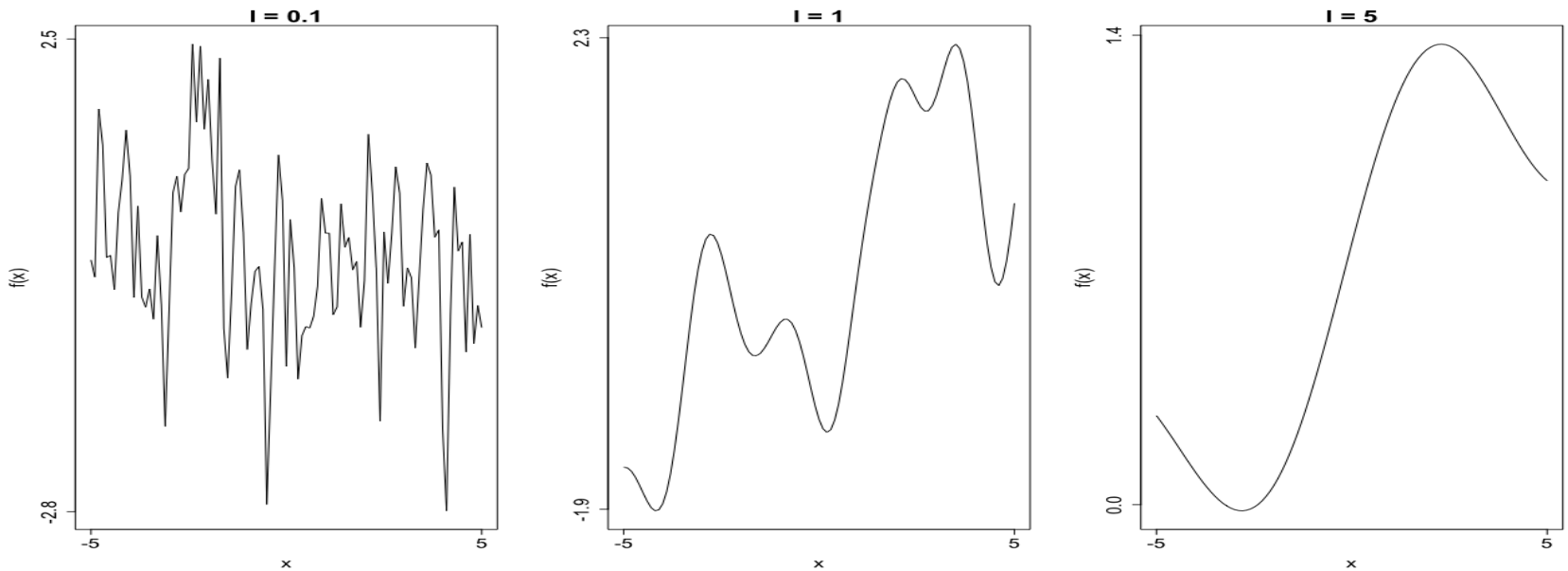


Figure: Three draws from the periodic covariance function.

Covariance Functions

The covariance function we will use is the square exponential

$$k(x, x'; \alpha, l) = \alpha^2 \exp\left\{ -\frac{(x - x')^2}{l^2} \right\}$$



Inference for Epidemics with GPs

How can we use GPs for modelling the infection rate?

1. Adopt a Bayesian Framework
2. Put a GP prior on infection rate
3. Use data augmentation to overcome intractability
4. Develop efficient MCMC algorithm to explore posterior density

Table of Contents

Building the Model

We will model the spread of a disease where the infection rate is distance dependent.

The individuals will be fixed on a 2D plane.

An example of this is the spread of Foot and Mouth Disease or Avian Flu.

Building the Model

We will model the spread of a disease where the infection rate is distance dependent.

The individuals will be fixed on a 2D plane.

An example of this is the spread of Foot and Mouth Disease or Avian Flu.

For example, we could assume i and j, $\beta_{i,j}$, is

$$\beta_{i,j} = \beta_0 \exp\{-\rho(i,j)\},$$

where $\rho(i,j)$ is the distance between i and j.

Animation

The Statistical Model

The likelihood function for this model is given by

$$\pi(\mathbf{i}, \mathbf{r}|\boldsymbol{\beta}, \lambda, \gamma) \propto \exp\Big(-\sum_{j=1}^{n}\sum_{k=1}^{N} \beta_{j,k}\big((r_j \wedge i_k) - (i_j \wedge i_k)\big)\Big) \prod_{j=1}^{n}\Big(\sum_{k\in\mathcal{Y}_j} \beta_{k,j}\Big)$$
$$\times \frac{\gamma^{n\lambda}}{\Gamma(\lambda)^n} \prod_{j=1}^{n}(r_j - i_j)^{\alpha-1} \exp\Big\{ -\gamma \sum_{j=1}^{n}(r_j - i_j)\Big\}.$$

We put the following GP prior distribution on β

$$\beta_{j,k} = \exp\big\{g(\rho(j,k))\big\}, \quad g \sim \mathcal{GP}(\boldsymbol{\mu}, \Sigma).$$

We also put an exponential prior distribution on γ, the infectious period parameter.

$$\gamma \sim \text{Exp}(\nu).$$

The Statistical Model

The posterior density is given by

$$\pi(\boldsymbol{\beta}, \gamma | \mathbf{i}, \mathbf{r}, \lambda) \propto \mathcal{GP}(\mathbf{g}) \exp\Big(- \sum_{j=1}^{n} \sum_{k=1}^{N} \beta_{j,k} \big((r_j \wedge i_k) - (i_j \wedge i_k) \big) \Big)$$
$$\times \prod_{j=1}^{n} \Big(\sum_{k \in \mathcal{Y}_j} \beta_{k,j} \Big) \gamma^{n\lambda} \exp\Big\{ - \gamma \sum_{j=1}^{n} (r_j - i_j) \Big\}$$
$$\times \gamma \exp\{-\nu\gamma\}.$$

We explore the posterior density using MCMC

- Data Augmentation for infection times
- MH for $\boldsymbol{\beta}$ and infection times
- Gibbs sampler for γ

Example

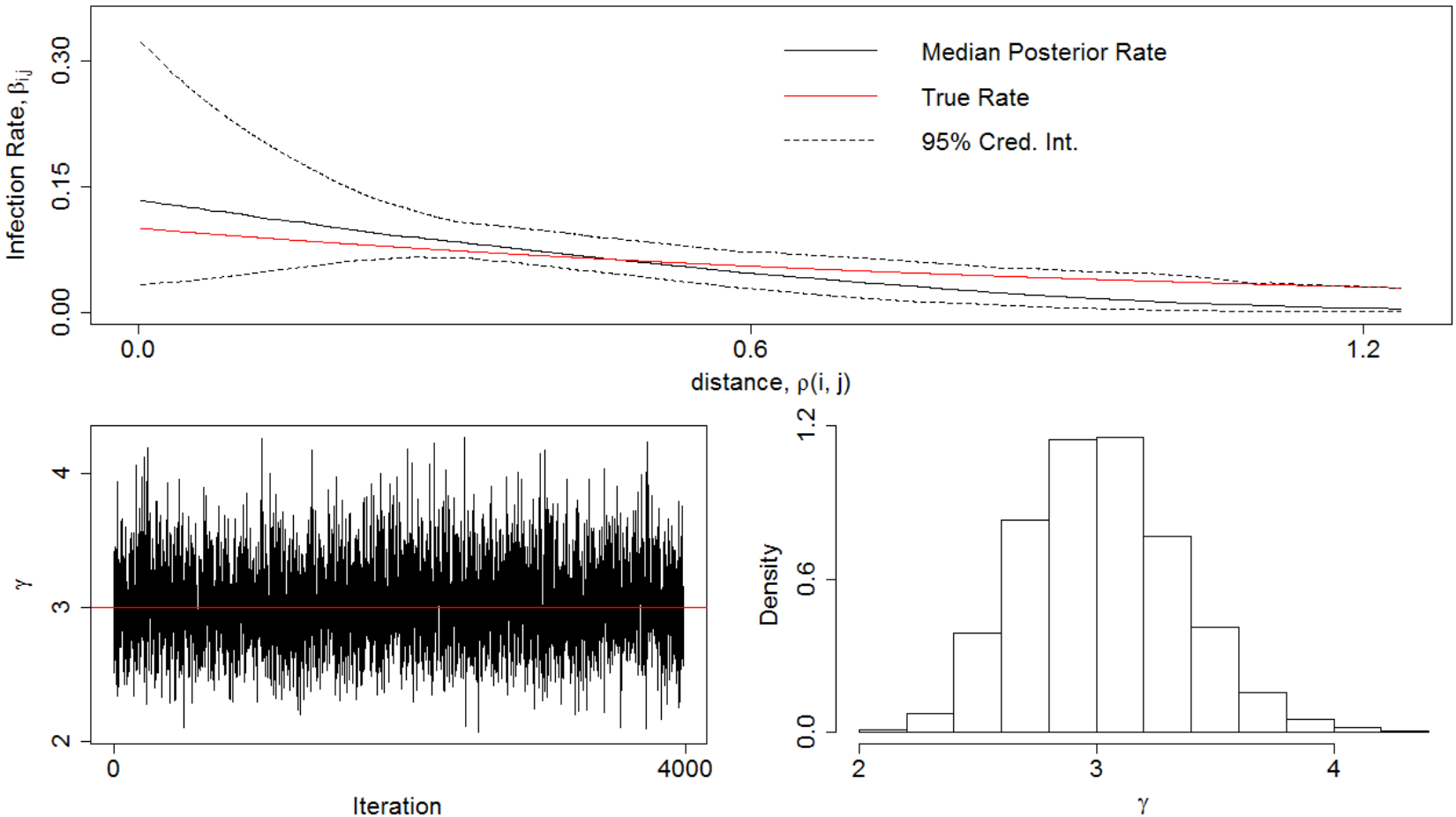


Figure: MCMC output when infection times are known.

Example

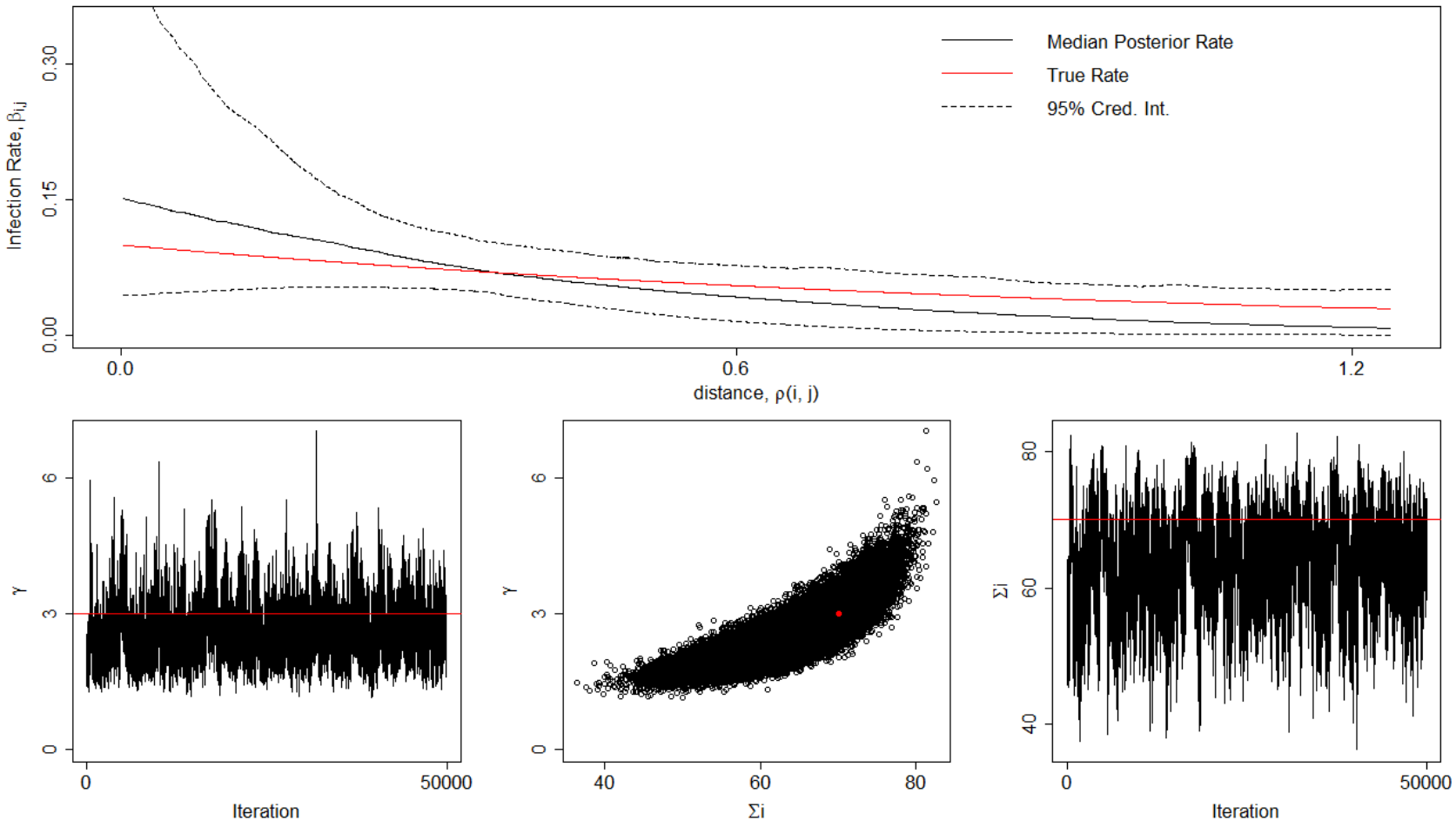


Figure: MCMC output when infection times are unknown.

Table of Contents

Including More Covariates

So far we have only looked at including distance in our model and we want to include other multiple inputs in our model.

For example, in the case of avian flu we might want to include:

- Distance
- Number of birds on each farm
- Type of farm
- Type of bird on each farm
- Farms which are owned by the same group

Efficiency

There are two main obstacles with MCMC methods with GPs:
The Covariance Matrix:

- Inverting and decomposing are computationally expensive.
- Can we choose strategic input points?
- How can we implement sparsification methods?

MCMC:

- MCMC can be very slow and computationally intensive.
- We're looking to make the code more efficient.

References