Stochastic CRNs

Lecture 3
SDS Estimation

Grzegorz Rempala
The Ohio State University

Statystyka Matematyczna XLV
Będlewo (POL)
Dec 5th 2019

1 Introduction

• We want to draw inferences on the transmission parameter $\lambda(\theta)$ and the distribution of the infectious period I by means of Maximum Likelihood (ML) theory.

• Consider a vector of counting processes $N = (N_1, \ldots, N_k)$, where each component $N_i(t)$ is (NHPP) counting the number of times a specific event has occurred up to time t.

• The probability that such an event occurs in $(t, t+h]$ given the history of the whole vector process up until t, denoted $\mathcal{H}_t$ satisfies

$$P(N_i(t+h) - N_i(t) = 1|\mathcal{H}_t) = h\lambda_j(t) + o(h), \quad j = 1, \ldots, k,$$
$$P(N(t+h) - N(t) = 0|\mathcal{H}_t) = 1 - h\lambda_0(t) + o(h).$$

where $\lambda_0(t) = \sum_i \lambda_i(t)$. Assume $\lambda_i(t) = \lambda_i(\theta_i; x(t))$ and let $\theta = (\theta_1, \ldots, \theta_k)$.

• Let $\mathbf{x} = \{x(t); 0 \le t \le T\}$ be an observed trajectory of $N(t); 0 \le t \le T$.

Complete data likelihood

• The likelihood corresponding to the i-th event (say of type ν_i) is just the joint density of the time and type of that event.

$$\lambda_0(\theta; x(t_{i-1}))\exp\{-\lambda_0(\theta; x(t_{i-1}))[t_i - t_{i-1}]\} \times \frac{\lambda_{\nu_i}(\theta_{\nu_i}; x(t_{i-1}))}{\lambda_0(\theta; x(t_{i-1}))}$$
$$= \exp\{-\lambda_0(\theta; x(t_{i-1}))[t_i - t_{i-1}]\}\lambda_{\nu_i}(\theta_{\nu_i}; x(t_{i-1}).$$

• The full likelihood is the product of these terms, together with a final term reflecting no event in the final interval $(t_n, T]$ giving the combined likelihood

$$\mathcal{L}(\theta; \mathbf{x}) = \prod_{i=1}^{n} \lambda_{\nu_i}(\theta_{\nu_i}; x(t_{i-1}))\exp\{-\lambda_0(\theta; x(t_{i-1}))[t_i - t_{i-1}]\}\exp\{-\lambda_0(\theta; x(t_n))[T - t_n]\}$$
$$= \prod_{i=1}^{n} \lambda_{\nu_i}(\theta_{\nu_i}; x(t_{i-1}))\exp\{-\sum_{i=1}^{k}\int_0^T \lambda_i(\theta_i; x(t))dt\}$$
$$= \exp\left\{\sum_{i=1}^{k}\left(\int_0^T \log(\lambda_i(\theta_i, x(t-)))dN_i(t) - \int_0^T \lambda_i(\theta_i; x(t-))dt\right)\right\}$$

Score equations and martinagles

• Here and elsewhere $dN_i(t)$ is a random variable equal to one at the jump of $N_i(t)$ and zero otherwise.

• We denote $\ell_T(\theta) = \log(\mathcal{L}(\theta; \mathbf{x}))$ and also write $x(t-)$ and $\lambda(t-)$ for the left limit of the data and intensity processes, resp.

• Maximum likelihood estimates may be obtained by differentiating the loglikelihood and solving the **score equations** for $k = 1, 2 \ldots$

$$\partial \ell_T(\theta)/\partial \theta_k = \partial_k \ell_T(\theta) = \sum_i \int_0^T \left(\frac{\partial_k \lambda_i(\theta; t-)}{\lambda_i(\theta; t-)} dN_i(t) - \partial_k \lambda_i(\theta; t-) dt \right) = 0.$$

• It follows that the log-likelihood (score) process $\{\ell_u(\theta_0); u \geq 0\}$, evaluated at the true parameter value θ_0 is a zero mean martingale. Note: since $\ell_0(\theta_0) \equiv 0$ then $E(\ell_u(\theta_0)) = 0$ for all $u \geq 0$.

Micro: Stochastic SIR (McKendrick 1926)

The classical biochemical reaction model at the level of unit (molecule)

- Three types of molecules: susceptibles (S), infectives (I), removed (R)
- Molecules combine after exponential holding time according to the current reaction rates (Gillespie algorithm)
- Rates given by *the law of mass action*

$$\begin{aligned} S + I &\xrightarrow{\lambda_1} 2I; \quad \lambda_1 = \beta SI \\ I &\xrightarrow{\lambda_2} R; \qquad \lambda_2 = \gamma I \end{aligned}$$

- The basic reproductive number $\mathcal{R}_0 = \beta/\gamma$.
- Let's look at this model more closely!

Micro: Stochastic SIR (II)

- $(S(t), I(t), R(t))$ state of the collection of units at time $t > 0$
- $S(0) = n$; $I(0) = m$; $R(0) = 0$;
- β overall transmission rate; γ recovery rate

Trajectory Equation

$$
\begin{aligned}
S(t) &= S(0) - Y_1\left(\frac{\beta}{n}\int_0^t S(u)I(u)du\right)\\
I(t) &= I(0) + Y_1\left(\frac{\beta}{n}\int_0^t S(u)I(u)du\right) - Y_2\left(\gamma\int_0^t I(u)du\right)\\
R(t) &= Y_2\left(\gamma\int_0^t I(u)du\right)
\end{aligned}
$$

- Y_1 and Y_2 are two independent unit Poisson processes.
- basic assumption: the population is *uniformly mixed*.

Micro: Stochastic SIR (III)

- Under our assumptions the micro SIR system is Markovian
- One may simulate its trajectory using *the Gillespie algorithm*
- adapted from stochastic biochemical networks literature

Gillespie Algorithm

1. Initiate at $(S(0), I(0), R(0))$
2. Assume you have the process value $(S(t), I(t), R(t))$ at $t \geq 0$
3. Calculate rates $\lambda_1(t) = \beta S(t) I(t)/n$ and $\lambda_2(t) = \gamma\, I(t)$
4. Set next transition time Δt as $Exp(\lambda_1(t) + \lambda_2(t))$
5. Set transition type (1 or 2) as $Ber\left(\frac{\lambda_1(t)}{\lambda_1(t)+\lambda_2(t)}\right)$
6. Update $(S(t'), I(t'), R(t'))$ at $t' = t + \Delta t$ and go to 2

Micro: Sellke Construction (I)

Another way of looking at stoch. SIR model from the perspective of an individual unit (Sellke 1983):

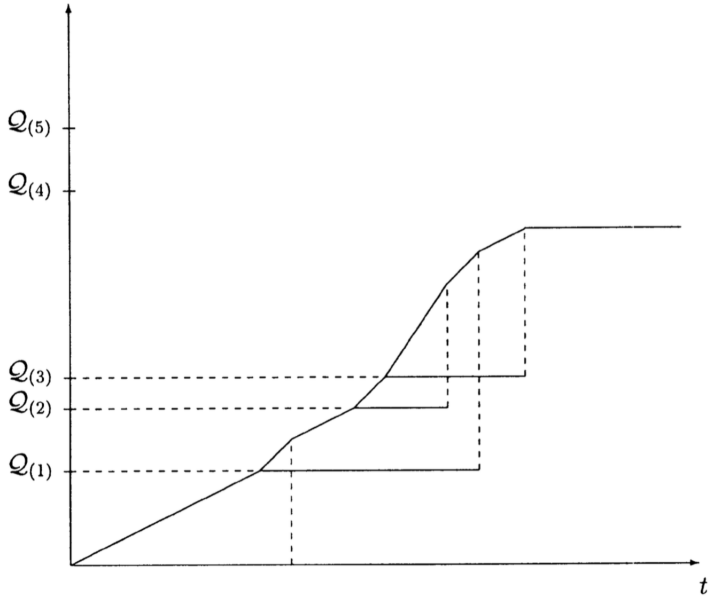
- Initial infectives are $-(m-1), -(m-2), \ldots, 0$ and initial susceptibles are $1, 2, \ldots, n$.
- Let $\mathcal{I}_{-(m-1)}, \mathcal{I}_{-(m-2)}, \ldots, \mathcal{I}_n \sim Exp(\gamma)$ be independent removal times
- Let $\mathcal{Q}_1, \mathcal{Q}_2, \ldots, \mathcal{Q}_n$ be $\sim Exp(1)$ be individual independent infection thresholds
- Let $I(t)$ be the number of infectives at time t, and set

$$A(t) = \frac{\beta}{n} \int_0^t I(u)du$$

to be the cumulative *infection pressure* or *hazard* exerted on a given susceptible up to time t.

Micro: Sellke Construction (II)

- The susceptible labeled i becomes infected when $A(t)$ exceeds $\mathcal{Q}_i$.
- The j-th infected susceptible remains infectious for a time $Exp(\gamma)$ and is then removed.



$m = 1, n = 5$

Micro: Sellke Construction (III)

- Let $\mathcal{Q} \sim Exp(1)$ be the threshold for random individual (unit)
- Let T_I be its *failure (infection) time*
- If $\mathcal{H}_t$ is history of the infections/recoveries up to t then

$$P(\mathcal{Q} > A(t)|\mathcal{H}_t) = P(\mathcal{Q} > A(t)|A(t)) = e^{-A(t)} = P(T_I > t|A(t))$$

Hence for a random unit failure time (conditionally on $A(t)$) is

$$P(T_I < t|A(t)) = 1 - e^{-A(t)}$$

Note: SC simply samples failure (infection) time T_I from the survival distribution with the cumulative hazard $A(t)$.

Macro: Kermack and Mc-Kendrick SIR (1927)

According to our LLN, as $n \to \infty$, assuming that

$$|(S(0), I(0), R(0))/n - (1, \rho, 0)| \xrightarrow{P} 0$$

the stochastic vector

$$\sup_{0<t<T} |(S(t), I(t), R(t))/n - (s_t, \iota_t, r_t)| \xrightarrow{P} 0$$

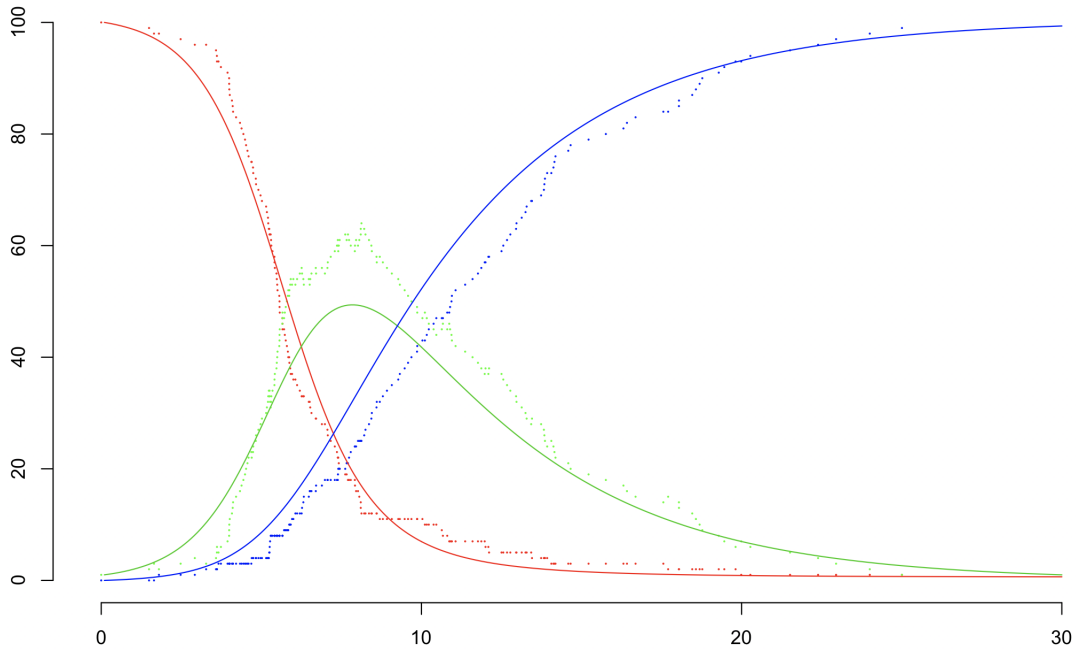
where (s_t, ι_t, r_t) is the solution of

$$\begin{aligned} \dot{s}_t &= -\beta s_t \iota_t \\ \dot{\imath}_t &= \beta s_t \iota_t - \gamma \iota_t \\ \dot{r}_t &= \gamma \iota_t \end{aligned}$$

$$s_0 = 1;\ \iota_0 = \rho > 0;\ r_0 = 0;$$

Mic Discrete vs Mac Continuous

$$S + I \xrightarrow{\lambda_1} 2I; \quad \lambda_1 = \beta SI \quad (\beta = 0.5)$$
$$I \xrightarrow{\lambda_2} R; \quad \lambda_2 = \gamma I \quad (\gamma = 0.3)$$
$$\mathcal{R}_0 = \beta/\gamma > 1 \quad P(\text{major outbreak}) > 0$$



Macro: Continuous SIR (I)

In different notation (CSIR(θ)) $\theta = (\beta, \gamma, \rho, \tau)$

$$\begin{aligned}
\dot{S}_t &= -\beta I_t S_t \\
\dot{I}_t &= \beta I_t S_t - \gamma I_t \\
\dot{R}_t &= \gamma I_t \\
S_0 &= 1; I_0 = \rho; R_0 = 0;
\end{aligned}$$

Denote $\mu_0 := \mathcal{R}_0$ and observe

$$\begin{aligned}
S_t &= Exp[-\beta \int_0^t I_s ds] = Exp[-\mu_0 R_t] \\
I_t &= \rho e^{-\gamma t} - \int_0^t \dot{S}_u e^{-\gamma(t-u)} du \\
R_t &= \gamma \int_0^t I_s ds \\
S_0 &= 1; I_0 = \rho; R_0 = 0;
\end{aligned}$$

Macro: CSIR (II)

- Mass (probability) transfer model from $\mathbb{S}$ to $\mathbb{R}$ compartment

- The state of a random unit u at t is $u_t \in \{\mathbb{S}, \mathbb{I}, \mathbb{R}\}$ with $u_0 = \mathbb{S}$
- $S_t = Exp[-\mu_0 R_t]$ is an *improper survival function* with $S_\infty = 1 - \tau > 0$
- $\tau = R_\infty - \rho > 0$ satisfies the final size equation

$$1 - \tau = Exp[-\mu_0(\tau + \rho)]$$

- Here τ is the probability of transfer out of state $\mathbb{S}$ (ever)
- $\mu_0 R_t = \beta \int_0^t I_s ds$ is the *cumulative hazard function*
- βI_t is the *hazard function* !

Micro CSIR (I)

- Let T_I be the exposure (infection) time of random u. By the law of total probability

$$P(T_I > t) = S_t = \tau \tilde{S}_t + 1 - \tau$$

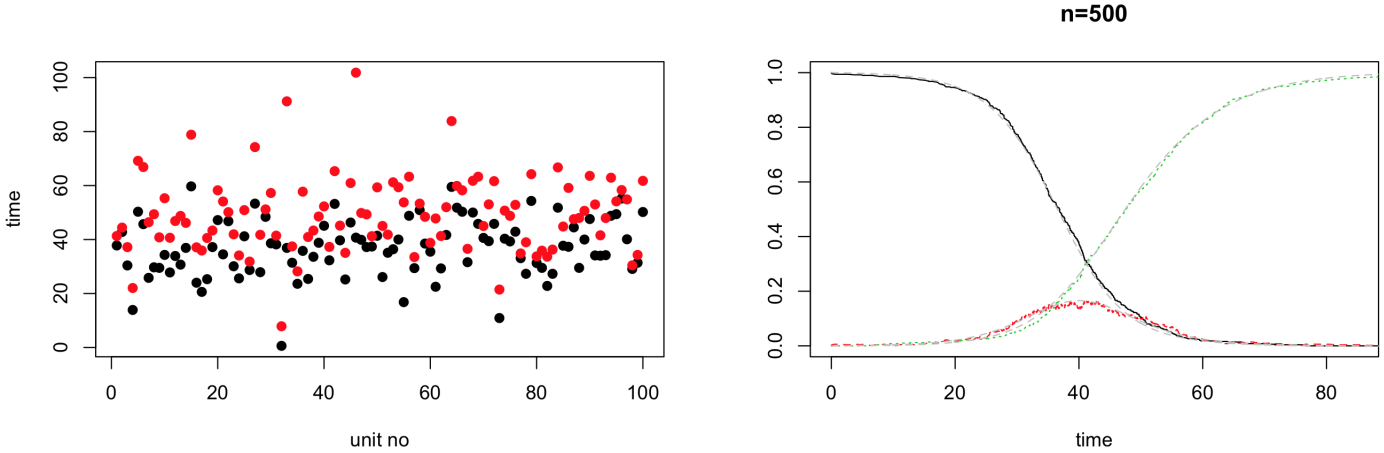
- $\tilde{S}_t = (S_t - 1 + \tau)/\tau$ is the proper conditional survival function
- Conditioned on u getting exposed (or failing, i.e. transferring out of $\mathbb{S}$)
- According to CSIR equation the density for $\tilde{S}_t$ is

$$f_\tau(t) = \frac{\beta}{\tau} I_t S_t = \frac{\beta}{\tau} I_t Exp[-\mu_0 R_t], \qquad t > 0$$

Micro CSIR: Distribution of Transfer Times

- Let T_R be the recovery time of random u with exposure time T_I.
- Consider joint distribution of (T_I, T_R). Note that for $s < t$ CSIR equation implies
$$P(T_I < s, T_R > t | \text{unit fails}) = \tilde{I}_s e^{-\gamma(t-s)}/\tau$$
with $\tilde{I}_s = I_s - \rho e^{-\gamma s}$
- Differentiating, we get
$$z(s,t) = \gamma e^{-\gamma(t-s)} f_\tau(s) \mathbb{1}[s < t]$$
- That implies T_I and $T_R - T_I$ are independent and $T_R - T_I \sim Exp(\gamma)$
- This allows us to simulate a random u from CSIR !

Micro CSIR: Empirical Trajectories



Left: sample of $n = 100$ histories of the units transferring from $\mathbb{S}$ to $\mathbb{R}$

Right: Macro CSIR vs empirical trajectories based on $n = 500$ unit histories

Itineraries in the symbolic dynamics theory (e.g., Hao and Bailin 1989)

Micro CSIR: Inference

- We have n sentinels monitored up to time t_{max} of which k failed at times $0 < t_1 < \ldots < t_k$ of which l recovered after $\delta t_1, \ldots, \delta t_l$.

- Since failure times and recovery times are all independent, the joint likelihood is then

$$\mathcal{L}(t_1, \ldots, t_k, \delta t_1, \ldots, \delta t_l, t_{max}) = \gamma^l \prod_{i=1}^{k} f_\tau(t_i) \prod_{j=1}^{l} e^{-\gamma \delta t_j} \, S_{t_{max}}^{n-k} \prod_{j=l+1}^{k} e^{-\gamma(t_{max} - t_j)}$$

where

$$f_\tau(t) = \beta \tau^{-1} I_t e^{-\mu_0 R_t}$$
$$S_t = e^{-\mu_0 R_t}$$

and $\tau = R_\infty - \rho$ satisfies the equation

$$1 - \tau = e^{-\mu_0(\tau + \rho)}$$

- MCMC: 2000 iterations with $n = 500$; all params within 5% error range.

1.1 Example: H1N1 Epidemic

In practice we would often impute recovery times, population size and/or use a continuous likelihood instead.

Hybrid Gibbs Sampler

1. Initiate $\theta = (\beta, \gamma, \rho, \kappa, \mu)$ from the prior distribution and set $n = k$.
2. Perform M-H step for the target conditional distribution of $(\theta|n)$ using the SDS likelihood
3. Calculate τ based on the current value of θ
4. Sample the conditional distribution of $(n|\theta)$ by drawing $n \sim NegBinom(k, \tau)$.
5. Return to step 2.
6. Repeat until convergence.

Binomial Random Measure

- Let $(E, \mathcal{E})$ be measurable sp. with a collection $\{X_i\}$ of iid rvs (stones) with law ν and
- Let K be a non-negative integer-valued, finite mean rv that is independent of $\{X_i\}$
- In all applications below we take $K \sim binomial(\tau, n)$
- Consider a random measure $\mathcal{N}$ associated with the pair (K, ν)

$$\mathcal{N}_\omega(A) = \mathcal{N}(\omega, A) = \sum_{i=1}^{K(\omega)} \mathbb{1}\!\mathrm{I}\,[A]\,(X_i(\omega)) \quad \text{for } \omega \in \Omega, \;\; A \in \mathcal{E}.$$

- On any test function $f \in \mathcal{E}_+$, $\mathcal{N}_\omega f = \sum_i^{K(\omega)} f(X_i(\omega))$.
- We write $\mathcal{N} f$, so that $\mathcal{N}(A) = \mathcal{N}\mathbb{1}\!\mathrm{I}\,[A]$.

SDS as BRM

- Consider the collection of n independent individuals (U_i) surveyed for symptoms of infectious disease as BRM $\mathcal{N} = (K, \nu \times Q)$ on the space $(E, \mathcal{E})$ where $E = \{(x, y) : 0 < x < y\}$

$$\nu(x) := f_\tau(x) = -\frac{\dot{s}_x}{\tau} \text{ and } Q(x, y) := H_x(y) \sim Exp(\gamma)\mathbb{1}\left[\{x < y\}\right](y)$$

- Each individual $U_i = (T_{i,I}, T_{i,R})$ is described by a pair of infection and recovery times
- Assume that at time $t > 0$ the collection of labels $L_t(U_i) \in \{S, I, R\}$ for $i = 1, \ldots, n$ is observed
- Notice that the law of BRM $\mathcal{N}$ is equal to the law of n independent trajectories simulated according to SDS model

CRM Partition

- Assume we have two observation points t_1 and t_2 with $0 < t_1 < t_2$
- The pair induces the partition of the state space:

$$\begin{aligned} E_{SS} &:= \{(x,y) \mid t_1 < t_2 < x < y\}, \quad E_{II} := \quad \{(x,y) \mid x < t_1 < t_2 < y\}, \\ E_{SI} &:= \{(x,y) \mid t_1 < x < y < t_2\}, \quad E_{IR} := \quad \{(x,y) \mid x < t_1 < y < t_2\}, \\ E_{SR} &:= \{(x,y) \mid t_1 < x < t_2 < y\}, \quad E_{RR} := \quad \{(x,y) \mid x < y < t_1 < t_2\}. \end{aligned}$$

- E.g., the observed labels of an individual in the partition E_{SS} are $\mathbb{S}$ at times t_1 and t_2, etc
- Define

$$\tilde{s}_t = (S_t - 1 + \tau)/\tau$$
$$\tilde{\iota}_t = I_t - \rho \exp\{-\gamma t\}$$

CRM Likelihood

- Setting $k = k_{SS}+k_{SI}+k_{SR}+k_{II}+k_{IR}+k_{RR}$, we can write the *CRM likelihood* as follows

$$\begin{aligned}
&P\left(\mathcal{N}(E_{SS}) = k_{SS}, \ldots, \mathcal{N}(E_{RR}) = k_{RR}\right) \\
&= \frac{n!}{k_{SS}! \cdots k_{RR}!(n-k)!} \tau^k (1-\tau)^{n-k} (\tilde{s}_{t_1} \tilde{s}_{t_2-t_1})^{k_{SS}} \\
&\quad \times \left(\tilde{s}_{t_1} \tilde{\iota}_{t_2-t_1}\right)^{k_{SI}} \left(\tilde{s}_{t_1}(1 - \tilde{s}_{t_2-t_1} - \tilde{\iota}_{t_2-t_1})\right)^{k_{SR}} \\
&\quad \times \left(\tilde{\iota}_{t_1} e^{-\gamma(t_2-t_1)}\right)^{k_{II}} \left(\tilde{\iota}_{t_1}(1 - e^{-\gamma(t_2-t_1)})\right)^{k_{IR}} \\
&\quad \times \left(1 - \tilde{s}_{t_1} - \tilde{\iota}_{t_1}\right)^{k_{RR}}
\end{aligned}$$

- We can marginalize out the unobservable counts retaining the multinomial likelihood form

Example: 2009 H1N1 at WSU

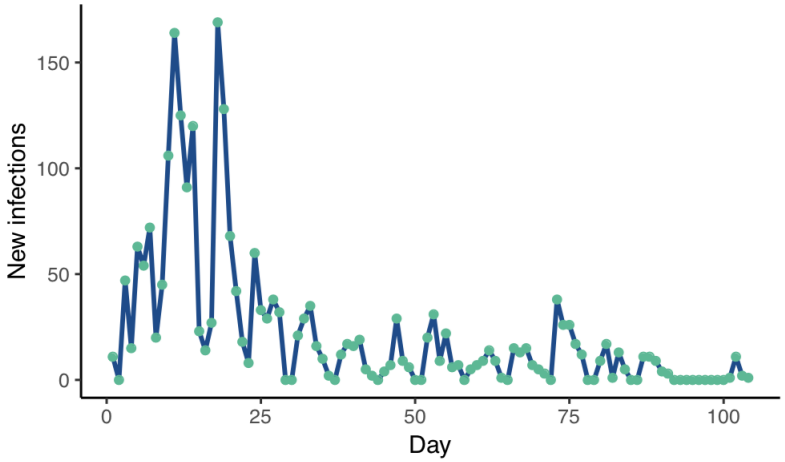


Figure 1: Daily new infections counts from WSU H1N1 outbreak.

Parameter	MAP	90% Credibility
n	7051	(6602, 7581)
β	0.1887	(0.185, 0.196)
ρ	0.0423	(0.04, 0.045)
$\mathcal{R}_0$	1.06	(1.04, 1.09)

Example: H1N1 Analysis Results

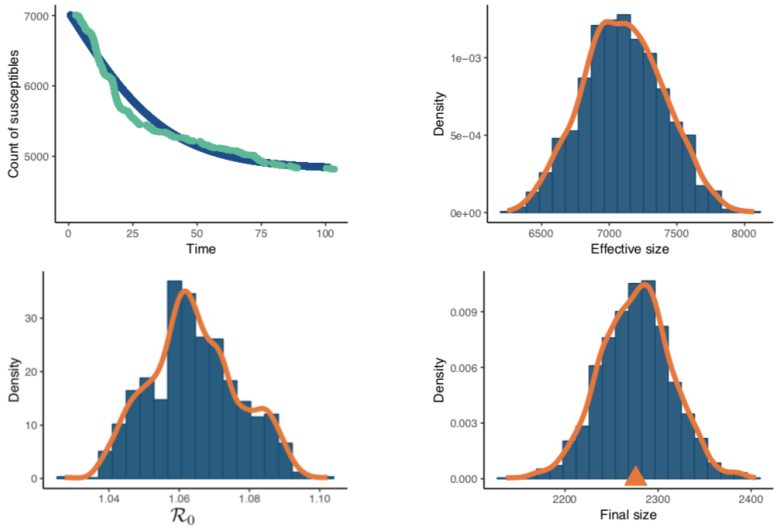


Figure 2: Top panel: fitted vs observed s_t curve and the posterior sample of the effective population size (n). Bottom panel: Posterior distributions of the basic reproduction number $\mathcal{R}_0$ and the final epidemic size. The latest distribution may be used to validate the model against actually observed data. The vertical line is drown at the actually observed epidemic size of 2276.

References

[1] Hakan Andersson and Tom Britton. *Stochastic epidemic models and their statistical analysis*, volume 151. Springer Science & Business Media, 2012.

[2] Boseung Choi, Grzegorz A Rempala, and Jae Kyoung Kim. Beyond the michaelis-menten equation: Accurate and efficient estimation of enzyme kinetic parameters. *Scientific reports*, 7(1):17018, 2017.

[3] Andrew Golightly and Darren J Wilkinson. Bayesian sequential inference for stochastic kinetic biochemical network models. *Journal of Computational Biology*, 13(3):838–851, 2006.

[4] Grzegorz A Rempala. Least squares estimation in stochastic biochemical networks. *Bulletin of mathematical biology*, 74(8):1938–1955, 2012.

[5] Grzegorz A Rempala, Kenneth S Ramos, and Ted Kalbfleisch. A stochastic model of gene transcription: an application to l1 retrotransposition events. *J Theor Biol*, 242(1):101–16, Sep 2006.

[6] Elissa J Schwartz, Boseung Choi, and Grzegorz A Rempala. Estimating epidemic parameters: Application to h1n1 pandemic data. *Mathematical biosciences*, 270:198–203, 2015.