MATH399: **Statistics Project**

Title: **Epidemic Modelling for Endemic Diseases**
Duration: 09/12-01/13

Name & ID: **Yihe Lu, 200838900**
Contact: sgylu3@liv.ac.uk

Supervisor: **Prof. D. Clancy**
Contact: dclancy@liverpool.ac.uk

Signature:
Date:

---

*My love is as a fever, longing still*
*For that which longer nurseth the disease,*
*Feeding on that which doth preserve the ill,*
*Th'uncertain sickly appetite to please.*

*My reason, the physician to my love,*
*Angry that his prescriptions are not kept,*
*Hath left me, and I, desperate, now approve*
*Desire is death, which physic did except.*

*Past cure I am, now reason is past care,*
*And frantic mad with evermore unrest;*
*My thoughts and my discourse as madmen's are,*
*At random from the truth vainly expressed:*
*For I have sworn thee fair, and thought thee bright,*
*Who art as black as hell, as dark as night.*

*-- Sonnet 147, William Shakespeare*

---

It is impossible for people to live without disease (or love).
May they not suffer too much. Live long and prosper.

I want to thank my supervisor, Prof. Clancy, who has helped me throughout the entire project.
I also thank my friends who spent time reviewing parts of my draft.

# Epidemic Modelling for Endemic Disease

Yihe Lu
*Department of Mathematical Sciences, The University of Liverpool, Liverpool, UK*

## *Abstract*

The SIRS model with demography is established for endemic diseases. With the equivalent deterministic model as an auxiliary tool, the stochastic model is studied by computer simulations. Comparison between the two transmission terms shows the frequency-dependent transmission stochastic model requires smaller population than the density-dependent transmission stochastic model for their equivalent deterministic models to be good approximations. Nåsell's approximations for quasi-stationary distribution are tested against simulation samples; the approximation for diseases like the common cold is fairly good, but the one for diseases like chickenpox is invalid.

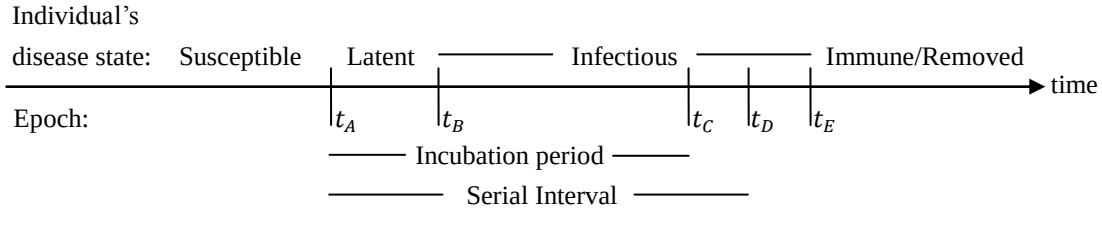
*keywords: Endemic disease; SIRS stochastic model with demography; Computer simulation; Transmission term; Quasi-stationary distribution*

## *1. Introduction*

Infectious diseases have contributed significantly to evolution, of which their historical impacts on human beings are just examples, documented the most recently and studied the most thoroughly (Mollison, 1995). However, it has been at most three centuries since the mathematical approach emerges in epidemic research (Bernoulli, 1760, cited in Daley and Gani, 1999).

Regardless of their various species, pathogens of infectious diseases are fairly small; they directly interact with cells, by which, however, they can affect the behavior of the physical systems of their host, and through the individual level, they could have noticeable influence on the population, locally or even globally (Mollison, 1995). The spread of pathogens at all these scales is a random process; whether some individual catches the disease, whether the immune system extinguishes the invading pathogens, and whether the disease spreads throughout the population or dies out, all occur at random. Nonetheless, the mathematical study of epidemics began with a deterministic model at the individual level (Bernoulli, 1760). Even though in *The Selfish Gene* (2006), a masterpiece explaining the essence

of all life forms by genes, Richard Dawkins, who firmly holds a gene-centred view, admits that it is more convenient to describe the behaviour of 'survival machines' at the individual level. For a similar but more practical reason, epidemiologists mainly work on modelling the spread of disease among individuals, because by understanding the epidemic process, people could learn how to control or prevent it, which may reduce the threat and risk of the disease, e.g. the first paper by Daniel Bernoulli (1760) was to help reduce the death rate of smallpox, in order to increase the population of France (Daley and Gani, 1999).




$t_A$: Infection occurs
$t_B$: Latency to infectious transition
$t_C$: Symptoms appear
$t_D$: First transmission to another susceptible
$t_E$: Individual no longer infectious to susceptibles (recovery or removal)
*Note:* $t_D$ should lie in interval $(t_B, t_E)$, so $t_D > t_C$ (as shown) and $t_D < t_C$ are both possible.

**Figure 1** – Progress of a disease in an individual.

Individuals are classified by a series of disease states since pathogens have consequent effects on them in different epochs, as illustrated in **Figure 1** (Daley and Gani, 1999). However, there are fewer compartments employed in most epidemic models for practical reasons. An SI (Susceptible-Infective) model with its population consisting only of susceptible and infective hosts is conventionally called the simple epidemic. An individual is susceptible before $t_A$ when infection occurs; after that, it becomes and remains infective, assumed to be infectious with no chance of recovery or removal. The simple epidemic is essentially the same as the logistic model (Verhulst, 1838), but of the growth of susceptible population (Daley and Gani, 1999). An SIR (Susceptible-Infective-Removal) model has one more compartment, namely removal, than the SI model. A removal is an individual removed from the epidemic for reasons based on different assumptions, e.g. disease-related death, immunity, isolation. Kermack and McKendrick (1927) firstly created this model for a homogeneous population, whose results have constituted a benchmark for many epidemic models. Two subsequent papers (Kermack and McKendrick, 1932, 1933) allow demographic variations, imperfect immunity, etc. and try to explain the mechanisms for endemic diseases. There are also other compartments, e.g. E for exposed groups (Liu, et al., 1987; Li and Muldowney, 1995), which corresponds to hosts in the latent period between $t_A$ to $t_B$ (cf. Bailey, 1957), C for carriers who are showing no symptoms but

infectious, i.e. in the state between $t_B$ to $t_C$ (Daley and Gani, 1999), etc.; they are less frequently employed because they may work well for certain diseases, but relatively lack of generality. This paper is concentrated on the SIRS model, which, like the SIR model, consists of three groups but further assumes that hosts removed will become susceptible again. Certainly in this case, hosts of removal are hosts of recovery, not someone dying for the disease and reviving from death; in fact, the disease is assumed not to be lethal at all, that is, there is no disease-related death (cf. **Figure 2**). Explicit definitions and detailed explanations of the model are contained in *Section 3*.

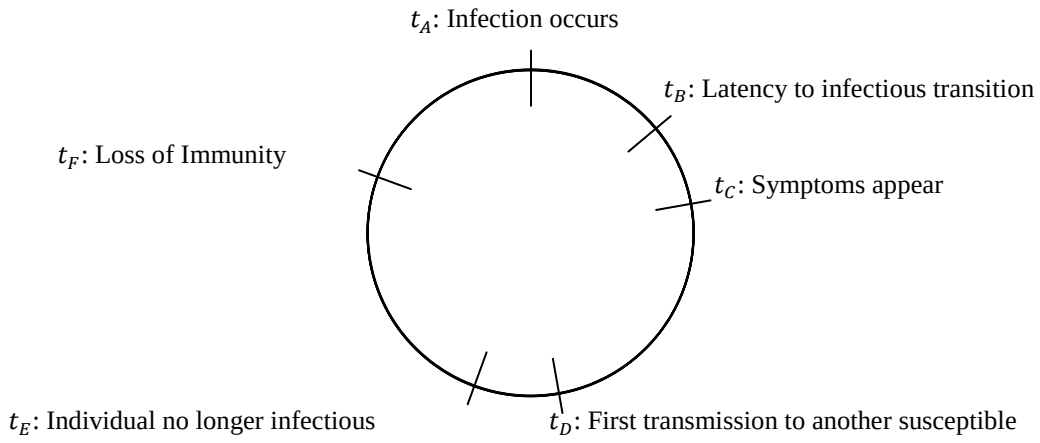



**Figure 2** – Progress of a disease in an individual who loses immunity and becomes susceptible again (cf. **Figure 1**).

Though all the previous models are studying epidemics which are essentially random processes, most of the works aforementioned (Bernoulli, 1760; Verhulst, 1838; Kermack and McKendrick, 1927, 1932, 1933; Liu, et al., 1987; Li and Muldowney, 1995) are concentrated on deterministic models. For models with a sufficiently large population, the mean numbers in the infection process would be characterised satisfactorily by the corresponding deterministic model. However, the deterministic model may approximate the process unfairly when the population size is small (Daley and Gani, 1999). Hence, although a deterministic model is simpler than its equivalent stochastic model, analysis of the stochastic model is to be preferred if possible (Andersson and Britton, 2000). Earliest stochastic models attribute to Mckendrick (1926) and Reed and Frost (1928); the later one was unpublished but more influential at that time (Daley and Gani, 1999; Andersson and Britton, 2000). Bartlett (1949) firstly introduced the lack-of-memory property, i.e. the Markovian property, to stochastic epidemic models, which effectively simplifies the mathematical analysis. The Markovian property is applied in this paper as well, because homogeneity is assumed and thus only numbers of hosts of different groups affect the epidemic process, but the process is independent of the history of any host or group, i.e. lack of memory (see *Section 2*).

This paper focuses on the SIRS stochastic model with demography. The model construction with the theory behind is to be explicitly demonstrated in *Section 2 & 3* and the discussion in *Section 4* is on the results of computer simulations. The comparison between the theoretical approximations and the simulation results of the stochastic model is to be the main purpose. Finally, conclusions are to be drawn in *Section 5*.

## *2. Continuous-time Markov chains*

### 2.1 Initial Definition

A continuous-time stochastic process $\{X(t), t \geq 0\}$ is defined as a continuous-time Markov chain if for all $s, t \geq 0$ and $i, j \in N^{*1}, x(u), 0 \leq u < s$,

$$P\{X(t+s) = j | X(s) = i, X(u) = x(u), 0 \leq u < s\} = P\{\mathrm{X}(t+s) = j | X(s) = i\}.$$

If this probability is independent of $s$, then the continuous-time Markov chain is said to have homogeneous transition probabilities (Ross, 2010).

### 2.2 Parallel Definition

It can be proved that the continuous-time Markov chain with homogeneous transition probabilities is equivalent to a stochastic process that each time it enters state $i$,

(i) the amount of time before any transition, denoted $T_i$, must satisfy $T_i \sim \mathrm{Exp}(\lambda_i)$, and

(ii) the probability of any transition from state $i$ to state $j$, denoted $P_{ij}$, must satisfy $P_{ii} = 0, \sum_j P_{ij} = 1$ for all $i$,

where $P_{ij} = \frac{\lambda_{ij}}{\lambda_i}$ and $\lambda_i = \sum_j \lambda_{ij}$ with $\lambda_{ij}$ denoting the transition rate from state $i$ to state $j$ (Norris, 1997).

### 2.3 Stationary Distribution

In the study of Markov chains, stationary[2] distributions are useful in explaining properties in long run. That is,

$$P_j = \lim_{t \to \infty} P_{ij}(t) > 0$$

is independent of $i$, if the limit exists. An equivalent condition for a distribution

[1] $N^*$ is the set of nonnegative integers.

[2] The terms invariant and equilibrium mean the same.

$\pi = (P_1 \quad P_2 \quad \dots \quad P_j \quad \dots), j \in N^+$[3] to be stationary is

$$\pi Q = 0$$

where $Q$ is the Q-matrix of the continuous-time Markov chain, defined as $Q = (q_{ij}: i, j \in N^+)$ by setting $q_{ii} = -\lambda_i$ and $q_{ij} = \lambda_{ij}$ for $i \neq j$.

Obviously, there is always a trivial solution, $\pi = 0$. Other than the trivial solution, $Q$ has a stationary distribution $\pi$ which is unique up to scalar multiples, if $Q$ is irreducible and recurrent (Norris, 1997).

### 2.4 Computer Simulation

Both definitions are applicable in computer simulation. To simulate the stochastic process with the initial definition, every iteration should be based on a small time period, in which the transition probabilities are calculated and whether the occurrence of the transition is decided at random, but the parallel definition allows each iteration to be based on a transition, where the holding time and the event are randomly decided.

Hence, simulation with the parallel definition takes less running time as simulation with the initial definition spends much running time on iterations where no transition occurs. The parallel definition is thereby employed in all computer simulations.

### 2.5 Error Analysis

Essentially, for every different transition from state $i$ to state $j$, it has its own exponential distribution with the corresponding parameter $\lambda_{ij}$ and thus each time the stochastic process enters state $i$,

$$T_i = \min_j(T_{ij})$$

where $T_{ij}$ denotes the amount of time before a transition to state $j$.

A problem would raise when $T_{ij_1} = T_{ij_2} = T_i$ for two distinct states $j_1$ and $j_2$; intuitively, we have no idea which state would the process enters. The proof of Norris (1997) and the parallel definition are valid, because the probability of this coincidence, $p_c$, is 0, due to the time continuum of the process in theory.

In practice, however, a computer, thus the random generators, is essentially discrete, which means

$$p_c = P\{T_{ij_1} = T_{ij_2} = T_i | j_1 \neq j_2\} > 0$$

and the parallel definition might malfunction in this situation; as the definition ignores $p_c$ and assigns it to other events unspecified by and unknown to us, an error may occur somewhere.

Claim: $p_c$ is very small as a computer is very accurate.

[3] $N^+$ is the set of positive integers.

Proof: Suppose $T_{ij_1}$~Exp($\lambda_1$) and $T_{ij_2}$~Exp($\lambda_2$), that is, for $t \geq 0$,

$$T_{ij_1} = \lambda_1 e^{-\lambda_1 t}$$

$$T_{ij_2} = \lambda_2 e^{-\lambda_2 t}$$

Consider the smallest time step of computer algorithms, denoted by $\Delta t$, then

$$p_c = P\{T_{ij_1} = T_{ij_2} | j_1 \neq j_2\} = \sum_{n=0}^{\infty} \lambda_1 e^{-\lambda_1 n\Delta t}\Delta t \cdot \lambda_2 e^{-\lambda_2 n\Delta t}\Delta t$$

$$= \lambda_1\lambda_2\Delta t^2 \sum_{n=0}^{\infty} e^{-(\lambda_1+\lambda_2)n\Delta t} = \frac{\lambda_1\lambda_2\Delta t^2}{1-e^{-(\lambda_1+\lambda_2)\Delta t}}$$

Let $x = (\lambda_1 + \lambda_2)\Delta t > 0$, substitute

$$p_c = \lambda_1\lambda_2 \cdot \left(\frac{x}{\lambda_1+\lambda_2}\right)^2 \cdot \frac{1}{1-e^{-x}} = \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} \cdot \frac{x^2 e^x}{e^x - 1}$$

$$< \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} \cdot x^2 e^x \cdot \frac{1}{x} = \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} \cdot x e^x$$

$$= \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_2)^2} \cdot (\lambda_1+\lambda_2)\Delta t \cdot e^{(\lambda_1+\lambda_2)\Delta t}$$

$$= \alpha\Delta t \xi^{\Delta t}$$

where $\alpha = \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}$ and $\xi = e^{\lambda_1+\lambda_2}$, $\alpha, \xi > 0$.

Therefore,

$$\left|\frac{p_c}{\Delta t}\right| < \alpha \cdot \xi^{\Delta t}$$

Since $\Delta t$ is very small for any modern computer, $\xi^{\Delta t} \approx 1$.

$$\left|\frac{p_c}{\Delta t}\right| < \alpha$$

■

The proof of the claim shows that the error decays as the computer accuracy rises. For modern computers, therefore, we could ignore the error.

# *3. Epidemic modelling*

## 3.1 Assumptions

### *3.1.1 Fundamental assumption of homogeneity*

Homogeneity applies to the whole population, except for the disease state of individual hosts. Hence, only the numbers of hosts in different disease states and the parameters affecting them are studied.

### *3.1.2 Assumption of SIRS models with demography*

In these models, we assume there are 3 disease states: susceptible, infective and recovered. There are new hosts entering the system and natural death causing removal of hosts from it. Infection may occur when susceptible and infective hosts contact each other. Infective hosts will recover and then have immunity against the disease. After a period of time, the immunity vanishes, which means a recovered host becomes susceptible again.

*3.1.3 Deterministic versus stochastic models*

Deterministic and stochastic models are two main approaches in characterising an epidemic process. Deterministic models are usually regarded as the mean of stochastic ones. They approximate stochastic models quite well when the population is sufficiently large, but they may not be satisfactory for small populations (Daley and Gani, 1999). In this paper, the main discussion focuses on the results of simulations of stochastic processes, but all the following notations are always used as specified, in either deterministic or stochastic cases:

$S$: the number of susceptible hosts,

$I$: the number of infective hosts,

$R$: the number of recovered hosts with immunity ($R = 0$ for SI and SIS models),

$N$: the total number of hosts ($N = S + I + R$), and

$N_0$: the size of the entire population containing the concerning community.

$S, I, R$ and $N$ are variables dependent on time $t$, while $N_0$ is assumed constant because the time period of the disease spread is usually much shorter than the average lifespan of its hosts, which means there are not noticeable changes of the entire population before the disease dies out.

## 3.2 Parameters

Conventionally, most of the involved parameters are called rate, but in order to avoid confusion with the concept of rates of the stochastic process, parameter is preferred here. They are all defined as follows.

*3.2.1 $\beta$ - the transmission parameter:*

Since all epidemic models are concerned with disease, how the interaction between susceptible and infective hosts leads to successful infection is one of the core issues and has raised debates over its essence and form.

Begon et al. (2002) reviews the historical and modern terminologies and provides a clear generalisation of transmission terms, the number of new infective hosts, denoted $\Delta I$, is derived from

$$\Delta I = Scpv$$

where $c$ is conventionally called 'Force of Infection', which causes most of the debates, $p$ is the probability that a contact between susceptible and infective hosts occurs and $v$ is the rate that the contact does cause infection. In particular, we assume $v$ a constant for any certain disease and $p = \frac{I}{N}$ by the individual homogeneity.

The assumption for $c$ has two alternatives. The transmission is called frequency-dependent transmission (FT) if we assume

$$c = \eta$$

where $\eta$ is an arbitrary constant predetermined by the disease, or density-dependent transmission (DT) if we assume

$$c = \kappa N/A$$

where $\kappa$ is another arbitrary constant predetermined by the disease and $A$ is the area where the community live, which is often assumed constant.

By simple arithmetic, we obtain

$$\Delta I = \beta_f SI/N$$

where $\beta_f = \eta v$ for FT and

$$\Delta I = \beta_d SI/N_0$$

where $\beta_d = \kappa v N_0/A$ for DT.

In the rest of this paper, $\beta$ is used for both $\beta_f$ and $\beta_d$ for conventions and they are to be set at the same value for convenience. However, they should still be interpreted in two different ways; with their identical dimensions, the comparison between the FT and DT models is more meaningful (cf. Begon et al., 2002).

*3.2.2 $\mu$ - the immigration-death parameter:*

The community studied in this paper is not closed. We assume there are new susceptible hosts entering the community, i.e. immigration, and natural death of hosts. The assumption is useful as most diseases remain in the community for a period long enough to observe immigration and death. Notice the immigration and death rates are identically $\mu$, because $N_0$ is constant and thus these two rates should be the same. Another simplifying assumption in this paper is that there is no disease-related death; that is, the death rate of infective hosts is the same as the death rate of susceptible and recovered hosts.

*3.2.3 $\gamma$ - the recovery parameter:*

An infective host remains ill and infectious for some time and then recovers. $1/\gamma$ is the average infectious period, and if $\gamma \to 0$, the model becomes an SI model where once a host is infected, they remain infectious until their death.

*3.2.4 $f$ - the loss of immunity parameter:*

A recovered host will eventually lose immunity; the average immune period is $1/f$. If $f \to 0$, the model becomes an SIR model where each recovered host is immune against the disease, and if $f \to \infty$, the model becomes an SIS model where a host actually cannot get immunity, that is, once they recover, they become susceptible immediately.

*3.2.5 $R_0$ – basic reproduction ratio*

In epidemiology, $R_0$ can be generally explained as the number of hosts infected by one infective host throughout the entire infectious period in a purely susceptible population (Heffernan, et al., 2005). It immediately follows from the definition that when $R_0 < 1$, every infective hosts infects, on average, less than one susceptible host and thus the disease tend to die out, but when $R_0 > 1$, the disease tends to become endemic, spreading throughout the population. $R_0$ is thus a basic measurement of endemicity.

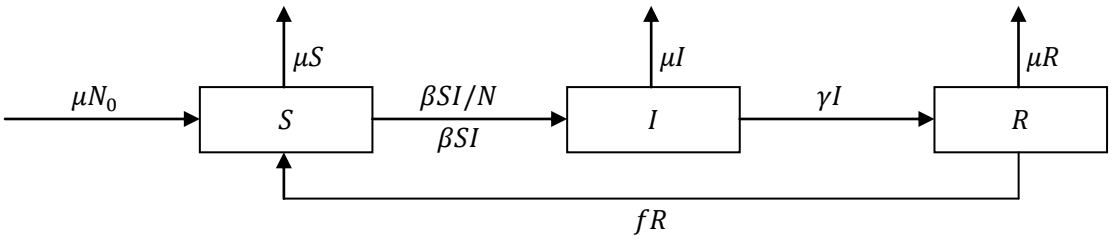



**Figure 3** – The SIRS model with demography.

## 3.3 Deterministic Model

*3.3.1 Formulation of the model*

**Figure 3** summarises the assumptions and definitions and an SIRS deterministic model can be formulated accordingly as a differential equation system:

$$\begin{cases} \dfrac{dS}{dt} = \mu N_0 - \mu S - \Delta I + fR, \\ \dfrac{dI}{dt} = \Delta I - \mu I - \gamma I, \\ \dfrac{dR}{dt} = \gamma I - \mu R - fR, \end{cases}$$

subject to the initial constraints specified by $S(0), I(0)$ and $R(0)$.

With the nonlinear term $\Delta I = \beta SI/N$ or $\beta SI/N_0$, an analytic solution to the system is not accessible, but the study of its equilibrium may still help us characterise the epidemic process.

*3.3.2 Equilibrium of the model*

First of all, the equilibrium population of the community can be solved. By definition,

$$\frac{dN}{dt} = \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = \mu N_0 - \mu S - \mu I - \mu R = \mu N_0 - \mu N$$

It is then straightforward to solve the equilibrium, by setting

$$\frac{dN}{dt} = 0$$

which immediately gives

$$\widehat{N} = N_0.$$

That is, the population will be settled at equilibrium, prefixed and independent of any variation of the community. Although the choice of FT or DT has no effect on the population, it may change the dynamics of the number of susceptible, infective and recovered hosts. To solve the equilibrium for these groups, set

$$\frac{dS}{dt} = \frac{dI}{dt} = \frac{dR}{dt} = 0$$

and solve for the DT model where $\Delta I = \beta SI/N_0$:

$$\begin{cases} \mu N_0 - \mu S - \beta SI/N_0 + fR = 0 \text{ ... (i)} \\ \beta SI/N_0 - \mu I - \gamma I = 0 \text{ ... (ii)} \\ \gamma I - \mu R - fR = 0 \text{ ... (iii)} \end{cases}$$

In equation (ii), if $I = 0$, then $R = 0$ by (iii) and thus $S = N_0$ by (i), which is exactly the trivial situation that the disease dies out; if $I \neq 0$,

$$\hat{S} = \frac{\mu + \gamma}{\beta} N_0.$$

Insert into (i) and (iii) and they become a linear system which has the unique solution:

$$\hat{I} = \frac{\mu + f}{\mu + f + \gamma}\left(1 - \frac{\mu + \gamma}{\beta}\right) N_0,$$

$$\hat{R} = \frac{\gamma}{\mu + f + \gamma}\left(1 - \frac{\mu + \gamma}{\beta}\right) N_0.$$

Similarly, we can solve in the case of FT where $\Delta I = \beta SI/N$. Notice $N = N_0$ at equilibrium, then the equation systems become exactly the same, that is,

$$\begin{cases} \hat{S} = \dfrac{\mu+\gamma}{\beta} N_0, \\ \hat{I} = \dfrac{\mu+f}{\mu+f+\gamma}\left(1-\dfrac{\mu+\gamma}{\beta}\right) N_0, \\ \hat{R} = \dfrac{\gamma}{\mu+f+\gamma}\left(1-\dfrac{\mu+\gamma}{\beta}\right) N_0. \end{cases}$$

Hence, the two models share the same equilibrium.

## 3.4 Stochastic Model

*3.4.1 Formulation of the model*

The stochastic epidemic model is of more interest than the deterministic one. It can be formulated as a Markov process with all possible transitions and their rates (see **Table 1**). The assumption of Markovian property is justified because all the transition rates are solely dependent on the current values of $S, I$ and $R$.

**Table 1** – the transition table of the SIRS model with demography

| Event | Transition from state $i$ to $j$ | Rate ($\lambda_{ij}$) |
|---|---|---|
| Immigration of susceptible | $(S,I,R) \to (S+1,I,R)$ | $\mu N_0$ |
| Death of susceptible | $(S,I,R) \to (S-1,I,R)$ | $\mu S$ |
| Death of infective | $(S,I,R) \to (S,I-1,R)$ | $\mu I$ |
| Death of recovered | $(S,I,R) \to (S,I,R-1)$ | $\mu R$ |
| Infection of susceptible | $(S,I,R) \to (S-1,I+1,R)$ | $\beta SI/N$ or $\beta SI/N_0$ |
| Recovery of infective | $(S,I,R) \to (S,I-1,R+1)$ | $\gamma I$ |
| Loss of immunity | $(S,I,R) \to (S+1,I,R-1)$ | $fR$ |

The parameter $\lambda_i$ can be calculated by summing up all the transition rates and the simulation of the epidemic process will then be obtain accordingly.

*3.4.2 Approach of the simulation*

In order to simulate the process, an exponentially distributed random generator: `random('exp', A)` is used frequently and essentially in Matlab programming (see *Appendix*), where `A` is the expectation of the exponential distribution.

With the claim in *Section 2.5* proved, it is unnecessary to generate $T_{ij}$ for all $j$ and then choose the smallest. We can confidently use the parallel definition, because it is robust for computer simulation.

For each iteration, i.e. at each state $i = (S, I, R)$, there are 3 steps:

(1) to decide the holding time: use the exponentially distributed random generator with its parameter A, taking value $1/\lambda_i$;

(2) to determine the thresholds: list $\lambda_{ij}$ in correspondence to all possible states $j$ in a fixed order and for each $j$, the threshold is obtained by adding up all rates of its preceding states and its own and dividing every sum through $\lambda_i$; and

(3) to choose a particular state $j$ to enter: use the uniformly distributed random generator: rand, which randomly produces value between 0 and 1, then compare the value to the thresholds, and choose to enter state $j$ if the value is smaller than the corresponding threshold but larger than the lower one.

Iterate all above steps and the simulation of the stochastic process is thus obtained and then plotted.

*3.4.3 Quasi-stationary distribution*

The stochastic process does not have any non-trivial stationary distribution, because clearly $I = 0$ is an absorbing state and all $I > 0$ are transient states, not recurrent, that is, every disease will eventually go extinct. However, the process might appear stable for a long time before the disease extinction. To describe this phenomenon, quasi-stationary distribution (QSD) is implicitly defined in Nåsell (2002), which approximates the model by a diffusion process and then estimates the mean values and standard deviations of the QSD by normal approximation.

Following the reparameterisation $R_0 = \frac{\beta}{\mu+\gamma}, \alpha_1 = \frac{\gamma+\mu}{\mu}$ and $\alpha_2 = \frac{\mu+f}{\mu}$, with $\alpha_1 \to \infty$, we obtain the approximation with means

$$\mu_S \approx \frac{1}{R_0} N_0,$$
$$\mu_I \approx \frac{\alpha_2(R_0 - 1)}{\alpha_1 R_0} N_0,$$
$$\mu_R \approx \frac{R_0 - 1}{R_0} N_0,$$

and the corresponding variances

$$\sigma_S^2 \approx \frac{\alpha_1}{\alpha_2 R_0^2} N_0,$$
$$\sigma_I^2 \approx \frac{R_0 - 1}{R_0^2} N_0,$$
$$\sigma_R^2 \approx \frac{\alpha_1}{\alpha_2 R_0^2} N_0$$

in the case $\alpha_2 = O(1)$, and another approximation with means

$$\mu_S \approx \frac{1}{R_0} N_0,$$
$$\mu_I \approx \frac{(R_0 - 1)\theta}{R_0(\theta + 1)} N_0,$$
$$\mu_R \approx \frac{R_0 - 1}{R_0(\theta + 1)} N_0,$$

and the corresponding variances

$$\sigma_S^2 \approx \frac{(R_0 + 1)\theta^2 + (R_0^2 + 2R_0)\theta + R_0}{R_0^2\theta(\theta + R_0)} N_0,$$
$$\sigma_I^2 \approx \frac{(R_0^2 - R_0 + 1)\theta^3 + (R_0^3 + R_0)\theta^2 + (R_0^3 + R_0^2 - R_0)\theta + R_0^2 - R_0}{R_0^2(\theta + 1)^2(\theta + R_0)} N_0,$$
$$\sigma_R^2 \approx \frac{(R_0^2 - R_0)\theta^3 + (R_0^3 - R_0 + 1)\theta^2 + (R_0^3 - 2R_0^2 + 3R_0)\theta + R_0}{R_0^2(\theta + 1)^2(\theta + R_0)} N_0$$

in the case $\alpha_2 = \theta\alpha_1$ where $\theta = O(1)$ (cf. Nåsell, 2002).

As the deterministic model suggests the same equilibrium for DT and FT, we might suppose that their QSDs have the same expectations and variances as well. However, it is not the interest of this paper to derive the results explicitly; this topic may leave for future research.

# *4. Simulation results*

## 4.1 Frequency-Dependent versus Density-Dependent Transmission

The deterministic model suggests that the FT and DT model would share the same equilibrium, but with the different transmission terms, they might approach different equilibria, or in different patterns. It would be interesting to study whether the stochastic process echoes the results. Since it is sophisticated to study the stochastic process analytically, computer simulations are employed to draw implicit conclusions.

The two systems seems to set at different equilibria eventually after $T = 0.5$ in **Figure 4**, which contradicts the results from the deterministic model. However, if we solve the deterministic equilibrium with the solutions derived in *Section 3.3.2*, we will get

$$\begin{cases} \hat{S} = \dfrac{\mu + \gamma}{\beta} N_0 = 104, \\ \hat{I} = \dfrac{\mu + f}{\mu + f + \gamma}\left(1 - \dfrac{\mu + \gamma}{\beta}\right) N_0 = 76.95, \\ \hat{R} = \dfrac{\gamma}{\mu + f + \gamma}\left(1 - \dfrac{\mu + \gamma}{\beta}\right) N_0 = 19.05 \end{cases}$$

which are respectively higher than $S(1), I(1)$ and $R(1)$ in **Figure 4**.

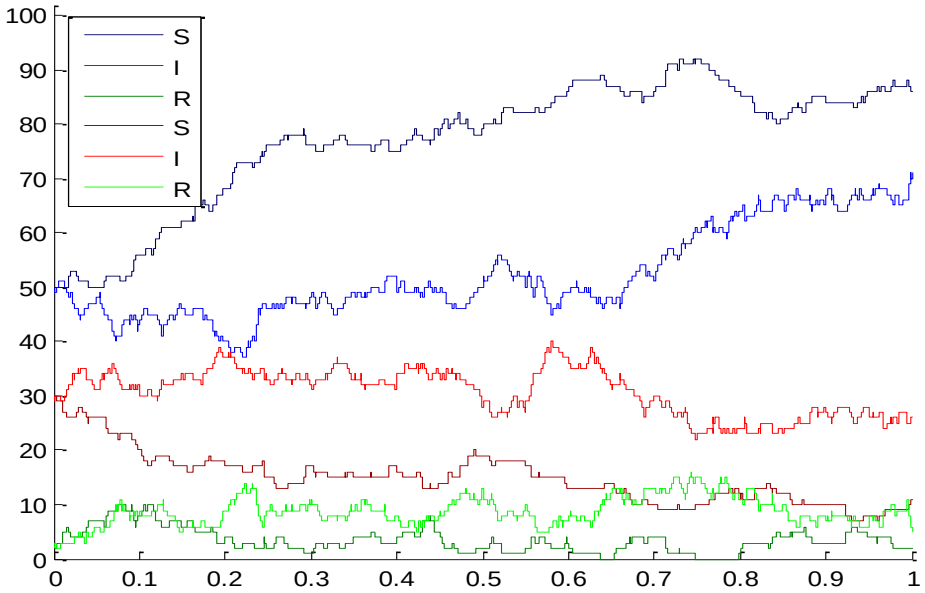



**Figure 4** – Simulation result with $R_0 = 1.92, \mu = 1/5, \beta = 10, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 1$; the darker lines are of DT model and the brighter ones are of FT model.

The phenomenon is caused by the small value of $\mu = 1/5$, making the population not at its equilibrium yet. A longer running time $T = 10$ (see **Figure 5**), comparing with the numerical solutions to the deterministic model, may thus be helpful.

The deterministic curves of the DT and FT models clearly illustrate different patterns, especially for $T < 2$ when the DT model goes against the direction of its equilibrium and then turns back around $T = 2$. For the stochastic models, we can see the disease somehow dies out with DT, in spite of $R_0 > 1$, and the extinction time is about $T = 2$ when the corresponding deterministic curve reaches its minimum; the randomness may thus lead to the extinction.

Though it becomes meaningless to compare and discuss the stochastic equilibria, we could conclude that the FT model seems to require a smaller sufficiently large population than the DT model does for the equivalent deterministic model to be its good approximation.

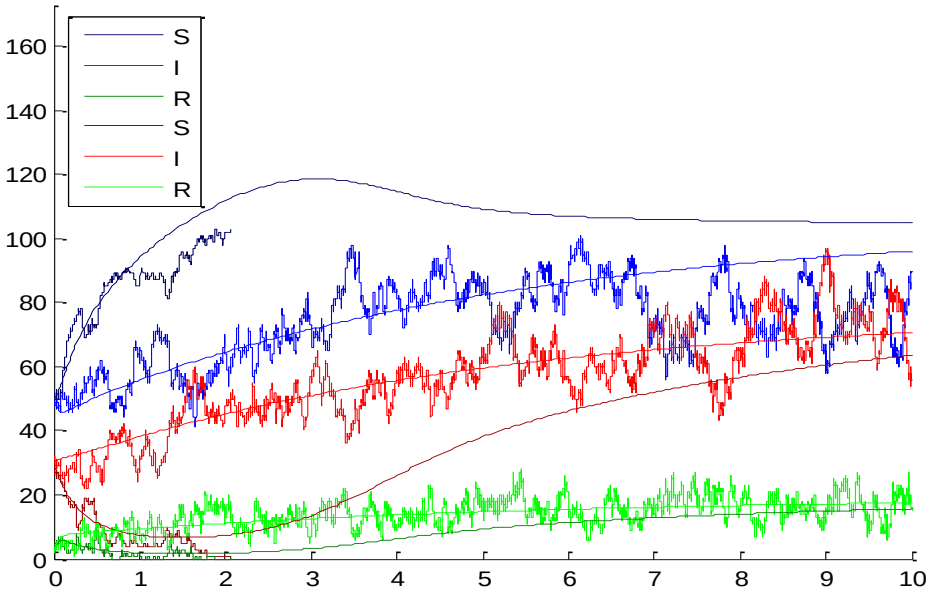



**Figure 5** – Simulation result with $R_0 = 1.92, \mu = 1/5, \beta = 10, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 10$; the curves are for the equivalent deterministic models, denoting the same groups of hosts with the same colours.

The two processes have distinct patterns at the beginning of the processes, because $N < N_0 \Rightarrow \beta SI/N > \beta SI/N_0$, i.e. FT is larger than DT, which means the disease is more infectious in the FT model. However, if the disease does not die out, as $N \to N_0$, the two transmission terms become indifferent, which allows the two processes then run seemingly in the same pattern (see **Figure 6**, where larger $\beta = 20$, increasing $R_0$ and thus preventing the disease from extinction).

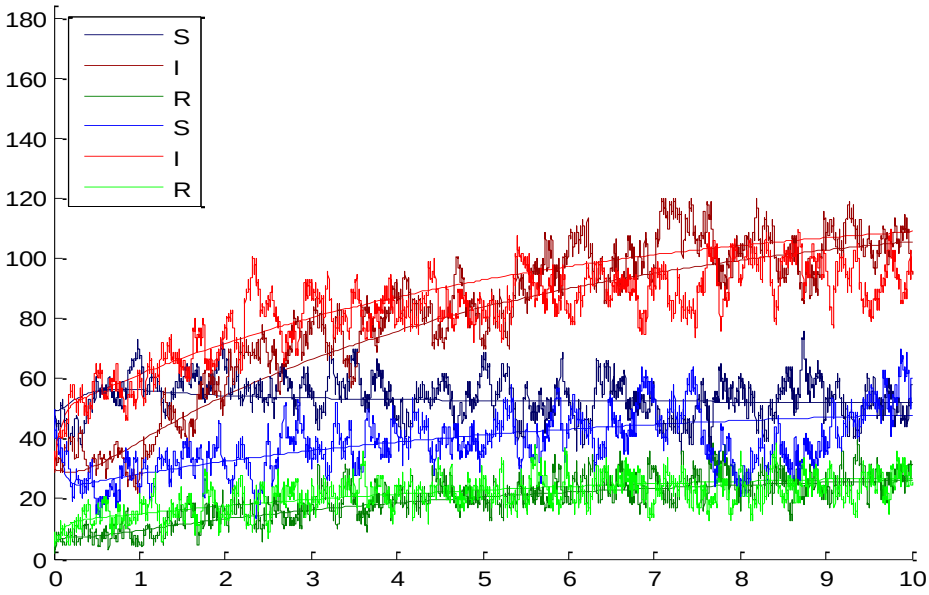



**Figure 6** – Simulation result with $R_0 = 3.85, \mu = 1/5, \beta = 20, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 10$.

Similarly, if the systems starts with $S(0) + I(0) + R(0) = N(0) = N_0$, there would not be noticeable differences between their patterns. However, if $S(0) + I(0) + R(0) = N(0) > N_0$, as in **Figure 7**, before $N \to N_0$, we have $N > N_0 \Rightarrow \beta SI/N < \beta SI/N_0$, which means the disease is now more infectious in the DT model, so there are more infective hosts in the system with DT.

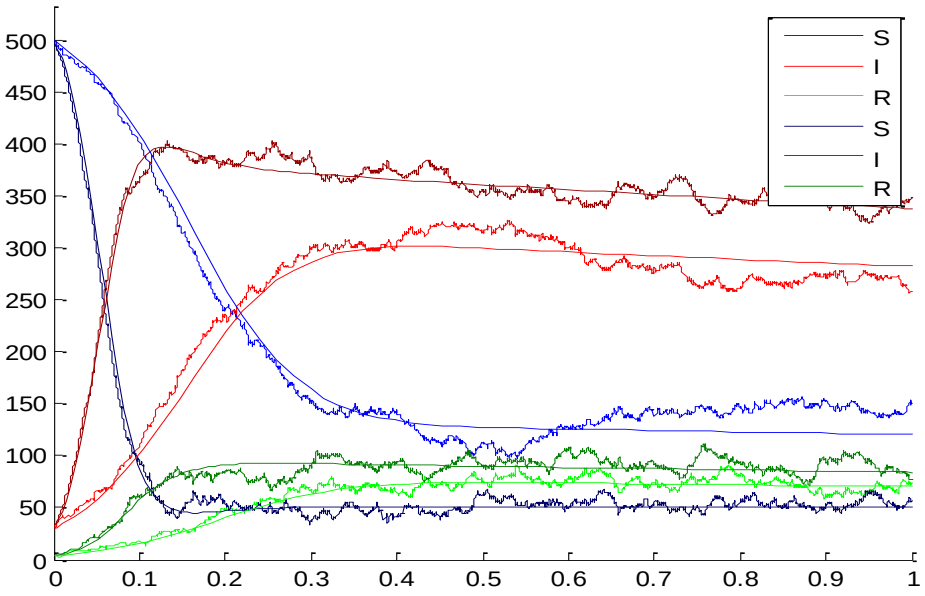



**Figure 7** – Simulation result with $R_0 = 3.85, \mu = 1/5, \beta = 20, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 500, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 1$.

With a long-time simulation (see **Figure 8**), the populations are decreasing to the equilibrium, and the differences between the two models decay as expected.

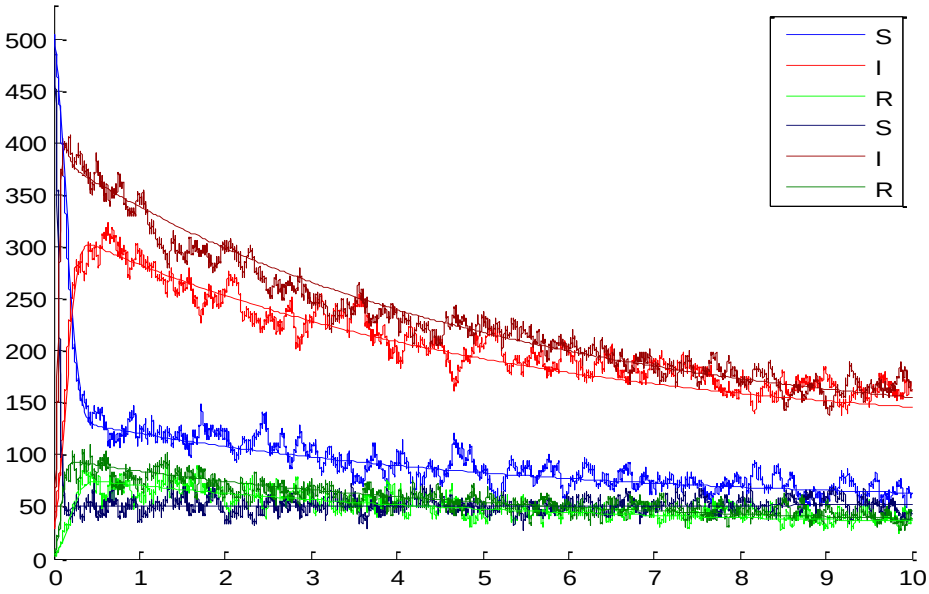



**Figure 8** – Simulation result with $R_0 = 3.85, \mu = 1/5, \beta = 20, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 500, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 10$.

Therefore, only if $N$ is very different from $N_0$, the two models yields two distinct patterns in the beginning of the processes, but if only the two processes have not ended in the short run, $N$ would be driven to $N_0$ and thus they appear to settle at the same equilibrium, as the deterministic equilibrium suggests.

## 4.2 Theoretical Approximations I: Case Study

### *4.2.1 Comparison with simulation sample*

The deterministic model estimates the mean values of the stochastic model better with a larger population, and further approximations of the expectations and the variances of the QSD are derived in Nåsell (2002). However, the theoretical results are not tested in a practical context. This section, therefore, is aimed at the justification of these theoretical approximations: the deterministic equilibrium and Nåsell's QSD approximations.

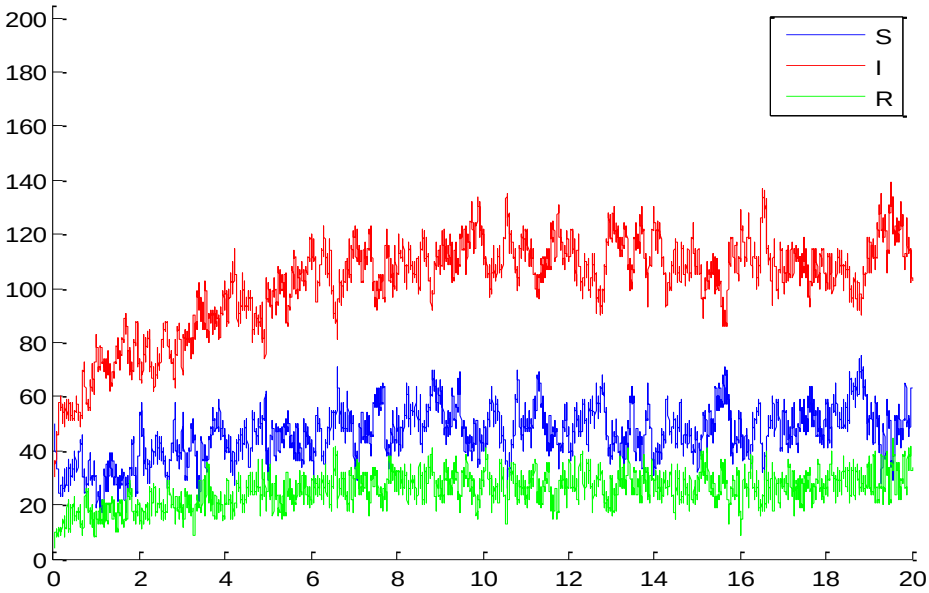



**Figure 9** – Simulation result with $\mu = 1/5, \beta = 20, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ for $T = 20$, which yields $R_0 = 3.85, \alpha_1 = 26, \alpha_2 = 101$.

The simulation illustrated in **Figure 9** is of the same case as in **Figure 4** but for the FT model only with a longer time $T = 20$, which seems to set about some equilibrium after $T = 10$. Indeed, the deterministic equilibrium is $\hat{S} = 52, \hat{I} = 118.63, \hat{R} = 29.37$, which seems to agree with **Figure 9**. Meanwhile, the QSD approximations are

$$\mu_S \approx 52.00, \sigma_S \approx 7.25,$$

$$\mu_I \approx 574.92, \sigma_I \approx 6.20,$$

$$\mu_R \approx 148, \sigma_R \approx 7.25,$$

which is dramatically different, and

$$\mu_S \approx 52.00, \sigma_S \approx 8.60,$$

$$\mu_I \approx 117.70, \sigma_I \approx 11.77,$$

$$\mu_R \approx 30.30, \sigma_R \approx 10.87,$$

which is much closer to the deterministic equilibrium and what is shown in **Figure 9**. The simulations for sample collection are all started at $S(0) = 52, I(0) = 119, R(0) = 29$, the round values of the deterministic equilibrium, because it is assumed close to the QSD and thus shortens the simulation time.

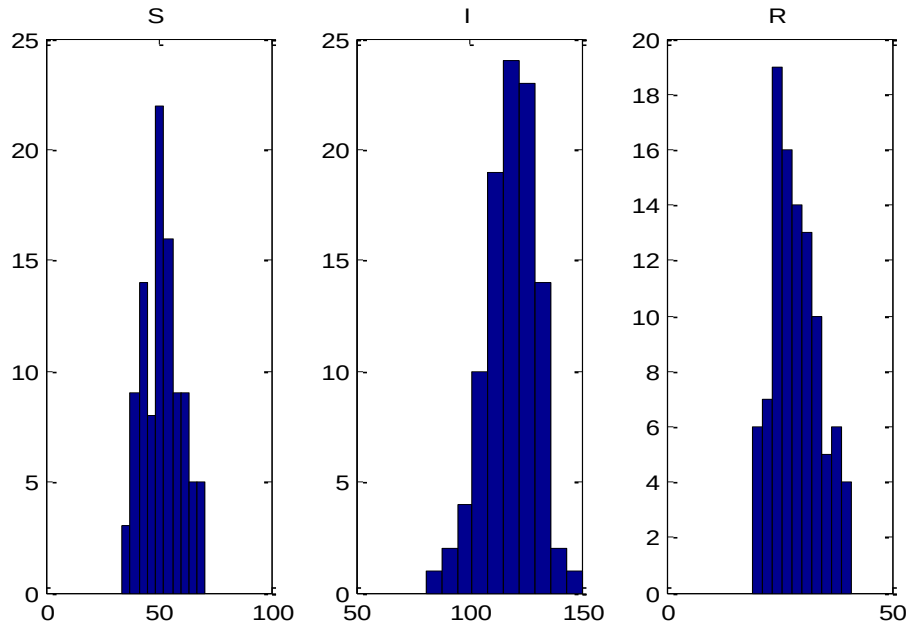



**Figure 10** – Histograms for 100 simulations with $\mu = 1/5, \beta = 20, \gamma = 5, f = 20, N_0 = 200$ and $S(0) = 52, I(0) = 119, R(0) = 29, N_0 = 200$ for $T = 5$, which yields $R_0 = 3.85, \alpha_1 = 26, \alpha_2 = 101$.

The sample gives **Figure 10**, which seems to justify the normal assumption, and

$$\bar{S}(5) = 51.80, \bar{I}(5) = 118.49, \bar{R}(5) = 28.86,$$

$$S_s(5) = 8.55, I_s(5) = 11.32, R_s(5) = 5.25.$$

The deterministic equilibrium and the second QSD approximation are close to the sample means and standard deviations, and we may conclude they are fair approximations.

*4.2.2 Statistical tests for approximations*

As the first QSD approximation is obviously invalid, we may draw conclusion by intuition. However, when the results seem close, proper statistical tests should be conducted to avoid heuristic errors.

For the deterministic equilibrium, the T-test is used to test the means with variances estimated by the sample:

H0: the sample means are equal to the deterministic equilibrium;

H1: the sample means are NOT equal to the deterministic equilibrium.

e.g. to test H0: $\bar{S} = \hat{S}$ against H1: $\bar{S} \neq \hat{S}$ with $\widehat{\sigma_S} = S_s$ as $T_S = \frac{\bar{S} - \hat{S}}{\widehat{\sigma_S}/\sqrt{n}} \sim t_{n-1}$.

For the QSD approximations, the Chi-squared test is used to test the variances and Z-test to test the means with variances estimated by the approximation:

H0: the sample variances are equal to the QSD approximation;

H1: the sample variances are NOT equal to the QSD approximation.

e.g. to test H0: $S_S^2 = \sigma_S^2$ against H1: $S_S^2 \neq \sigma_S^2$ as $X_S = \frac{(n-1)S_S^2}{\sigma_S^2} \sim \chi_{n-1}^2$.

H0: the sample means are equal to the QSD approximation;

H1: the sample means are NOT equal to the QSD approximation;

e.g. to test H0: $\bar{S} = \mu_S$ against H1: $\bar{S} \neq \mu_S$ with $\sigma_S$ as $Z_S = \frac{\bar{S}-\mu_S}{\sigma_S/\sqrt{n}} \sim N(0,1)$.

The similar notations and meanings apply for $I$ and $R$.

In the case in *Section 4.2.1*, these statistics can be obtained following above steps:

$$T_S = -0.23, T_I = -0.13, T_R = -0.96,$$

$$X_S = 97.89, X_I = 91.48, X_R = 23.05,$$

$$Z_S = -0.23, Z_I = 0.67, Z_R = -1.32,$$

while $t_{99}(0.95) = 1.66, \chi_{99}^2(0.90) = 117.41, \Phi(0.95) = 1.64$. That is, there are statistical evidences to support these results.

The first QSD approximation is invalid, because it should be under the condition $\alpha_1 \to \infty, \alpha_2 = O(1)$, while $\alpha_2 = 101$ is larger than $\alpha_1 = 26$. However, it is reasonable to treat both $\alpha_1$ and $\alpha_2$ sufficiently large, where $\theta \approx 4 = O(1)$, which makes the second QSD approximation work in this case.

## 4.3 Theoretical Approximations II: Nåsell's QSD

### *4.3.1 Reparameterisation*

In order to justify the accuracy and robustness of the Nåsell's QSD approximations, it would be more convenient to assign values to $R_0, \alpha_1$ and $\alpha_2$ first and recover the values of initial parameters $\beta, \gamma, f$ and $\mu$ (cf. *Section 3.4.3*).

Since $\beta, \gamma, f, \mu > 0 \Rightarrow R_0 = \frac{\beta}{\mu+\gamma} > 0, \alpha_1 = \frac{\gamma+\mu}{\mu} > 1, \alpha_2 = \frac{\mu+f}{\mu} > 1$, one can choose any values for the reparameterisation in order to agree with the conditions, and then with the information of $\mu$, by straightforward arithmetic,

$$\beta = R_0 \alpha_1 \mu,$$

$$\gamma = (\alpha_1 - 1)\mu,$$

$$f = (\alpha_2 - 1)\mu.$$

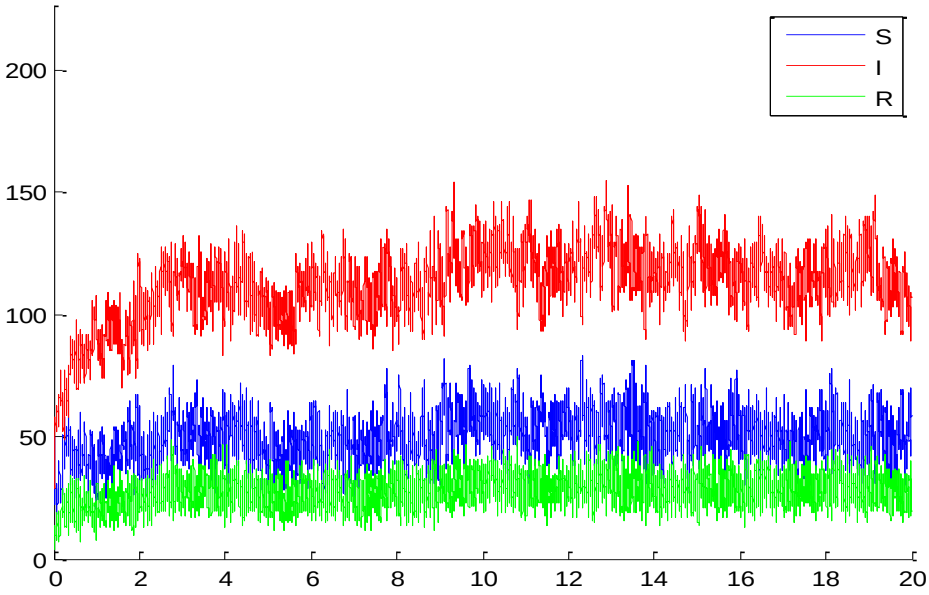



**Figure 11** – Simulation result with $R_0 = 3.85, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 100, \gamma = 25, f = 100$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 20$.

In order to keep the same values for the reparameterisation, all other parameters are scaled by $\mu = 1$. Even though the process in **Figure 11** seems to share the same equilibrium with **Figure 9**, their patterns appear to be distinct. Apparently, with larger parameters, the process approaches to the equilibrium more quickly. Essentially, larger parameters lead to larger rates, more events tend to occur, and less holding time a new transition would wait.

However, a 100-time sampling yields

$$\bar{S}(5) = 51.67, \bar{I}(5) = 117.86, \bar{R}(5) = 29.12,$$

$$S_s(5) = 8.31, I_s(5) = 10.38, R_s(5) = 5.50,$$

which are very close to the case when $\mu = 1/5$ and the histograms in **Figure 12** are similar to those in **Figure 10**.

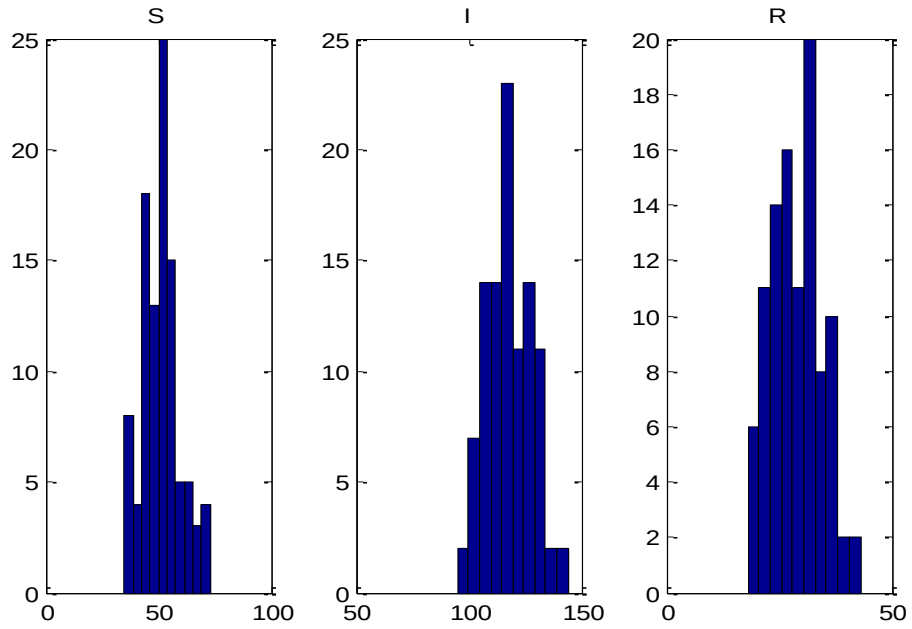



**Figure 12** – Histograms for 100 simulations with $\mathrm{R}_0 = 3.85, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 100, \gamma = 25, \mathrm{f} = 100$; $\mathrm{S}(0) = 52, \mathrm{I}(0) = 119, \mathrm{R}(0) = 29, \mathrm{N}_0 = 200$ and $\mathrm{T} = 5$.

Although **Figure 11** appears to have more 'randomness' in the process, it actually has more events (which occurs at random, though) and is essentially equivalent to the previous process with a longer running time. The reasoning can be straightforward: by specifying $R_0, \alpha_1$ and $\alpha_2$, the scaling of $\mu$ has the same effects on other parameters and the same effect on the total transition rate, which changes the holding time, but the chance for any particular event to happen stays indifferent if we know a transition is to occur.

Though, intuitively, there are differences between **Figure 9** and **Figure 11**, they are essentially the same up to time scaling. It can also be drawn from the fact that there are no terms containing $\mu$ in the deterministic equilibrium or the QSD approximations; that is, any legal value of $\mu$ would yield the same results. Therefore, it is more convenient and without any loss of generality to specify $\mu = 1$, which gives

$$\beta = R_0\alpha_1,$$

$$\gamma = \alpha_1 - 1,$$

$$f = \alpha_2 - 1.$$

Recall in either approximation in Nåsell (2002), $\alpha_1 \to \infty$, which means that $\alpha_1$ is large. It has been shown in the previous case that $\alpha_1 = 26, \alpha_2 = 101$ yields reasonable approximations under the condition $\alpha_2 = \theta\alpha_1$ where $\theta = \mathrm{O}(1)$, which implies for $\alpha_1 \to \infty$, it is unnecessary to take too large

a value. For consistence, the same values are assigned in cases under this condition.

*4.3.2 Common cold*

The common cold is a typical disease, from which a patient is easy to recover, but the immunity period is short, that is, it agrees with the condition $\alpha_1 \to \infty, \alpha_2 = \theta\alpha_1$ where $\theta = O(1)$. Due to the practical difficulty of taking 'infinity' as value, large $\alpha_1 = 26$ and $\alpha_2 = 101$ are taken and thus $\theta = 3.88$. For the mediocre $R_0 = 3.85$, the previous sections have proved the the QSD approximations are valid.

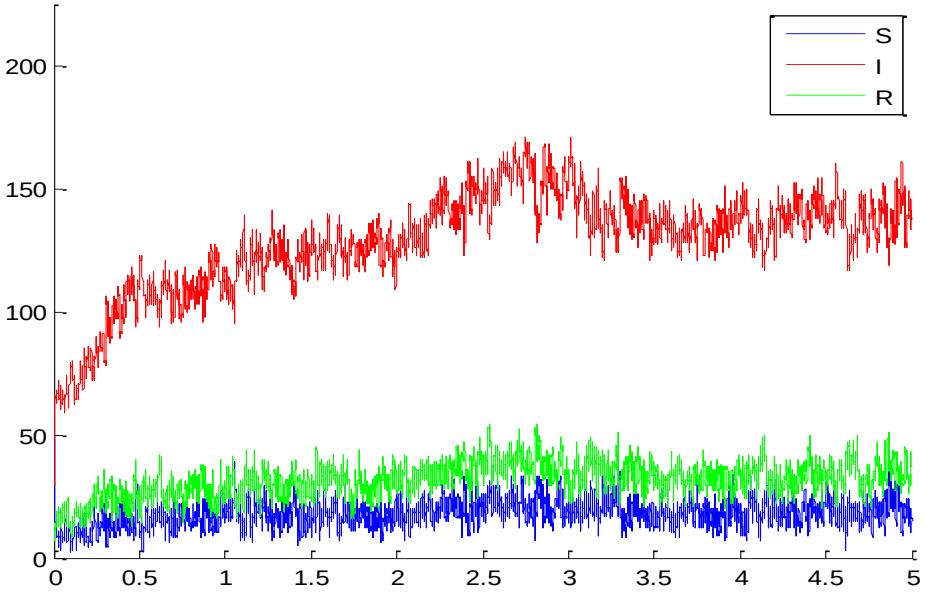



**Figure 13** – Simulation result with $R_0 = 10, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 260, \gamma = 25, f = 100$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 5$.

With $R_0 = 10$, the process seems to approach equilibrium when $T = 3$ in **Figure 13**. Apparently, the large $R_0$ causes the disease spreading throughout the population. Indeed, the QSD approximation:

$$\mu_S \approx 20.00, \sigma_S \approx 4.88,$$

$$\mu_I \approx 143.15, \sigma_I \approx 12.26,$$

$$\mu_R \approx 36.85, \sigma_R \approx 11.90,$$

suggest there are more infective hosts.

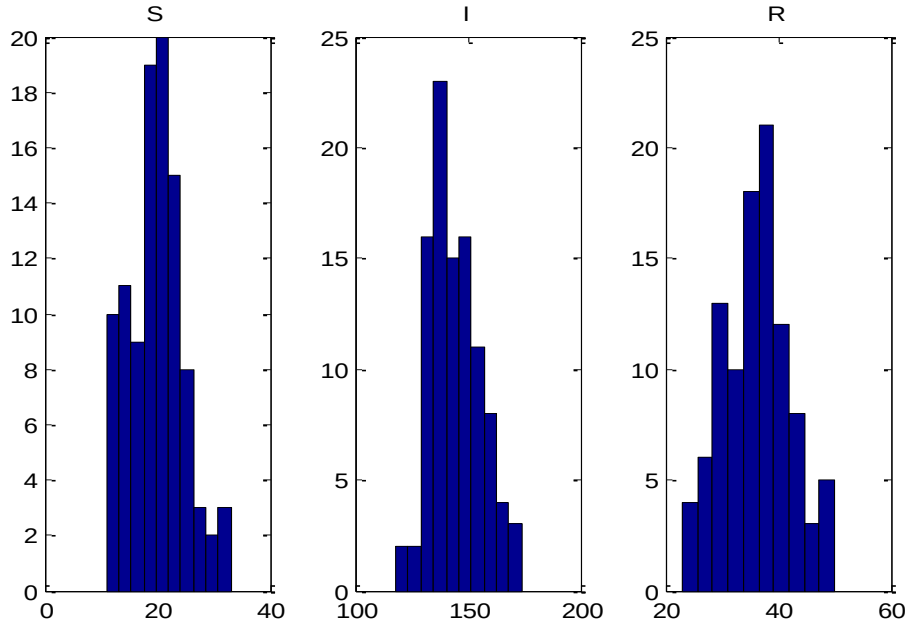



**Figure 14** – Histograms for 100 simulations with $R_0 = 10, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 260, \gamma = 25, f = 100$; $S(0) = 20, I(0) = 144, R(0) = 36, N_0 = 200$ and $T = 2$.

The larger value of $\mu$ practically shortens the running time. A sample of 100 simulations is summarised in **Figure 14**. Though the histograms for $S, I$ in seem skewed, the numerical results are close:

$$\bar{S}(5) = 19.92, \bar{I}(5) = 144.59, \bar{R}(5) = 36.28,$$

$$S_s(5) = 4.83, I_s(5) = 11.47, R_s(5) = 5.93.$$

Indeed, by proper statistical tests as in Section 4.2.2,

$$X_S = 96.99, X_I = 86.67, X_R = 24.54,$$

$$Z_S = -0.16, Z_I = 1.17, Z_R = -0.48,$$

which provides statistic evidences to accept the QSD approximation.

Actually, any large value of $R_0$ would yield similar pattern under this condition: the majority of the population are infected and the disease is endemic, and the QSD approximation works reasonably.

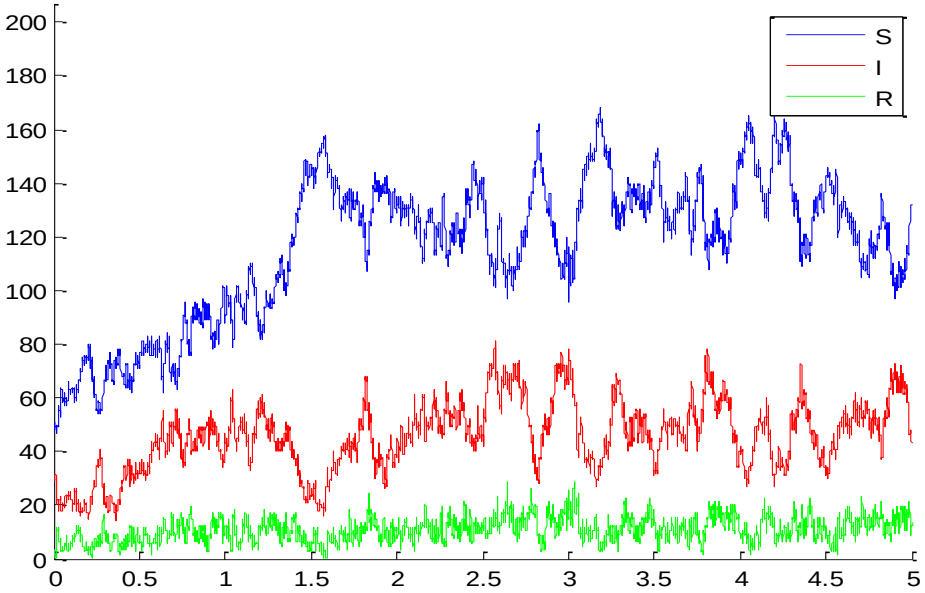



**Figure 15** – Simulation result with $R_0 = 1.5, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 39, \gamma = 25, f = 100$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 5$.

Since the QSD approximation only works for $R_0 > 1$, it might been expected to be worse with smaller $R_0$. **Figure 15** is depicted with $R_0 = 1.5$; meanwhile it also gives the QSD approximation:

$$\mu_S \approx 133.33, \sigma_S \approx 15.92,$$

$$\mu_I \approx 53.02, \sigma_I \approx 11.55,$$

$$\mu_R \approx 13.65, \sigma_R \approx 8.40.$$

A 100-simulation sample gives **Figure 16** and

$$\bar{S}(5) = 137.65, \bar{I}(5) = 51.45, \bar{R}(5) = 12.65,$$

$$S_s(5) = 16.37, I_s(5) = 11.90, R_s(5) = 4.21.$$

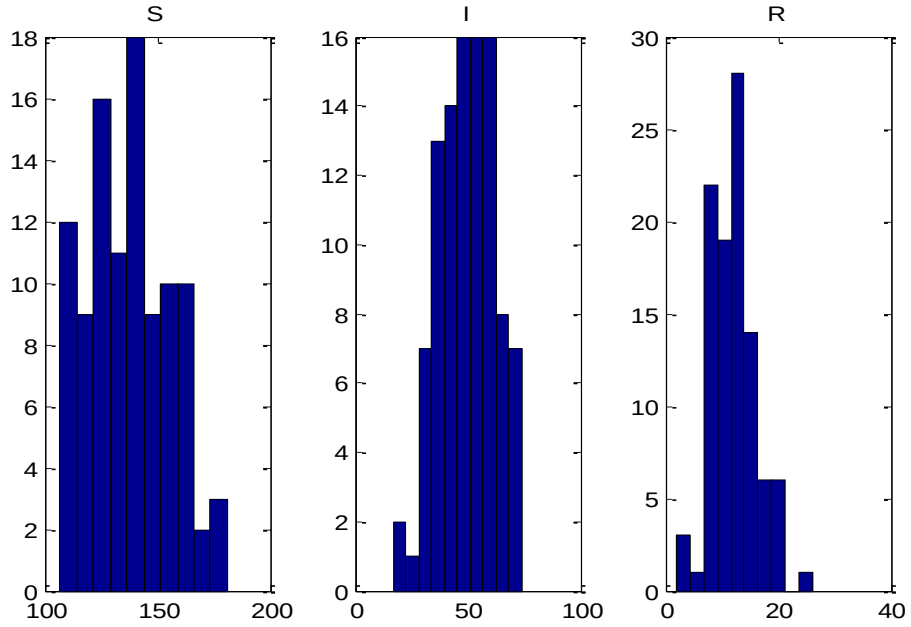



**Figure 16** – Histogram for 100 simulations with $R_0 = 1.5, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 39, \gamma = 25, f = 100$; $S(0) = 133, I(0) = 53, R(0) = 13, N_0 = 200$ and $T = 2$.

The results seem valid, since the numerical results still appear to agree with each other and the histograms tend to justify the normal assumption. However, they are all worse than the previous case:

$$X_S = 104.72, X_I = 105.10, X_R = 24.87,$$

$$Z_S = 2.71, Z_I = -1.35, Z_R = -1.19.$$

Although the approximation for variances is still acceptable as $\chi^2_{99}(0.90) = 117.41$, it estimates $S$ bad as $\Phi(0.995) = 2.56$.

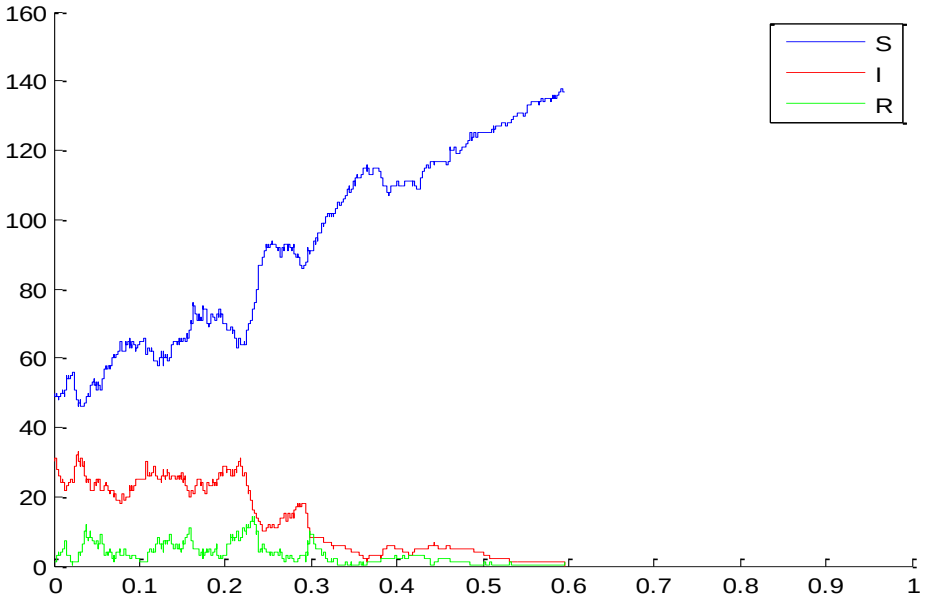



**Figure 17** – Simulation result with $R_0 = 1.1, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 28.6, \gamma = 25, f = 100$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 1$.

The approximation is expected to be worse when $R_0 \to 1$, which means the disease are much less infectious, i.e. it is very likely to die out early (see **Figure 17** where $R_0 = 1.1$). The QSD approximation:

$$\mu_S \approx 181.82, \sigma_S \approx 19.82,$$

$$\mu_I \approx 14.46, \sigma_I \approx 12.22,$$

$$\mu_R \approx 3.72, \sigma_R \approx 6.95$$

are not likely to be valid because it is not likely to obtain any equilibrium in this case, except for the trivial solution where the disease dies out (see **Figure 18**). Thus, the QSD approximation is invalid in this case.

Therefore, N åsell's QSD approximation works fairly well for diseases like the common cold.

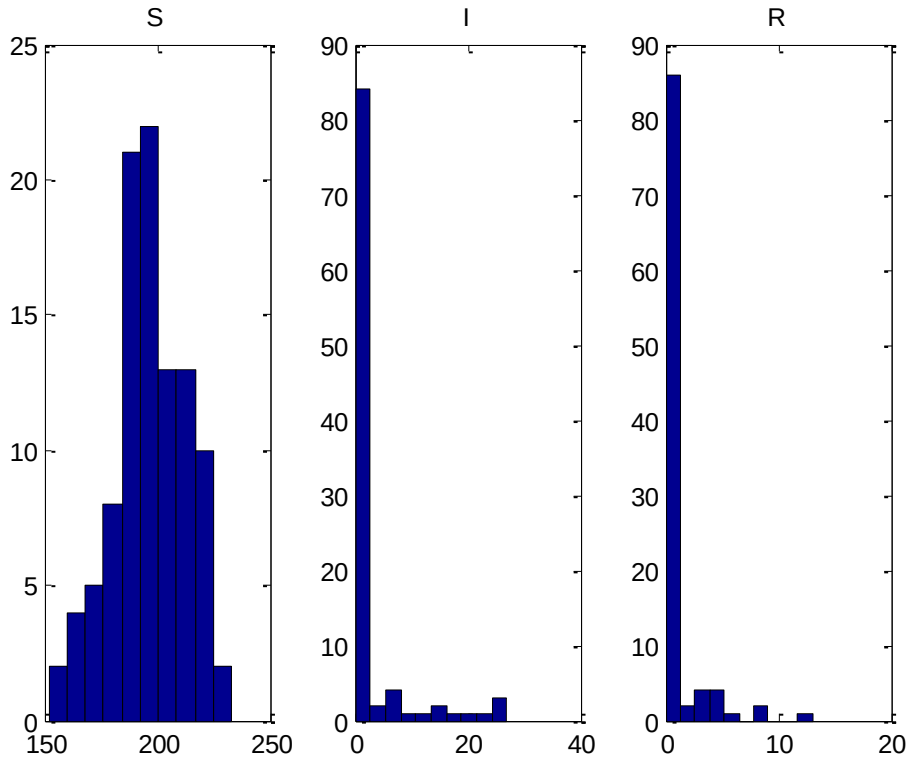



**Figure 18** – Histogram for 100 simulations with $R_0 = 1.1, \alpha_1 = 26, \alpha_2 = 101$, when $\mu = 1$, $\beta = 28.6, \gamma = 25, f = 100$; $S(0) = 182, I(0) = 15, R(0) = 4, N_0 = 200$ and $T = 2$.

*4.3.3 Chickenpox*

Chickenpox is a typical disease, from which a patient is easy to recover, and the immunity period is lifelong, that is, it agrees with the condition $\alpha_1 \to \infty, \alpha_2 = 0(1)$. As the loss of immunity takes long time, i.e. $\alpha_2 = f + 1$ in this case becomes small, fewer recovered hosts become susceptible again. Hence, it can be expected the processes are more likely to finish early than those under the previous conditions, that is, the disease dies out quickly.

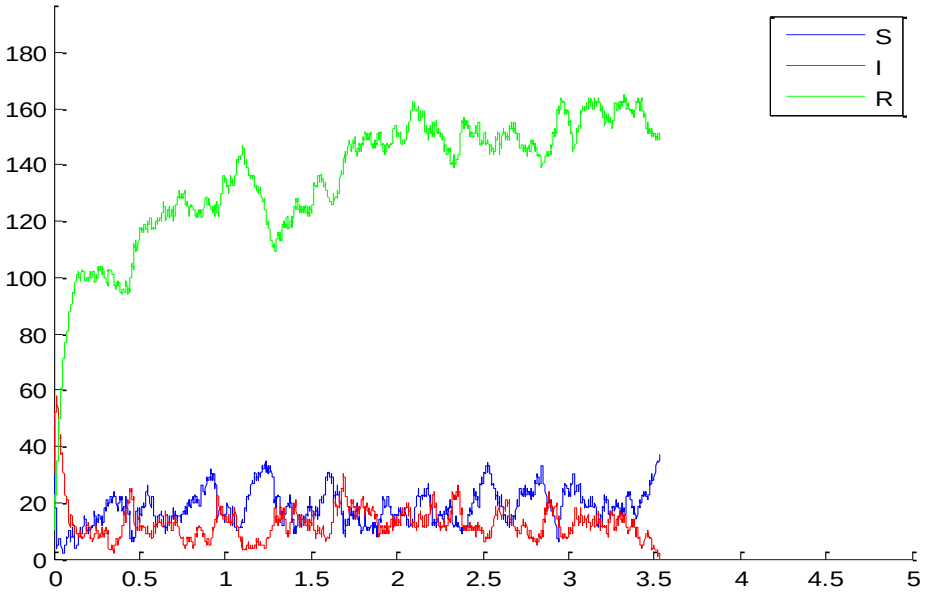



**Figure 19** – Simulation result with $R_0 = 10, \alpha_1 = 26, \alpha_2 = 2$, when $\mu = 1$, $\beta = 260, \gamma = 25, f = 1$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 5$.

The process, even with $R_0 = 10$, shows a distinct pattern in **Figure 19** from that in **Figure 13**, that the majority of the population are recovered hosts with immunity as expected. The pattern is different from that in **Figure 17** as most of the hosts are susceptible in that community and it appears, intuitively, that the process has a QSD from $T = 0.5$ to $T = 3$.

The QSD approximation:

$$\mu_S \approx 20.00, \sigma_S \approx 0.39,$$

$$\mu_I \approx 13.85, \sigma_I \approx 4.24,$$

$$\mu_R \approx 180, \sigma_R \approx 0.39,$$

seems to overestimate the value of $R$.

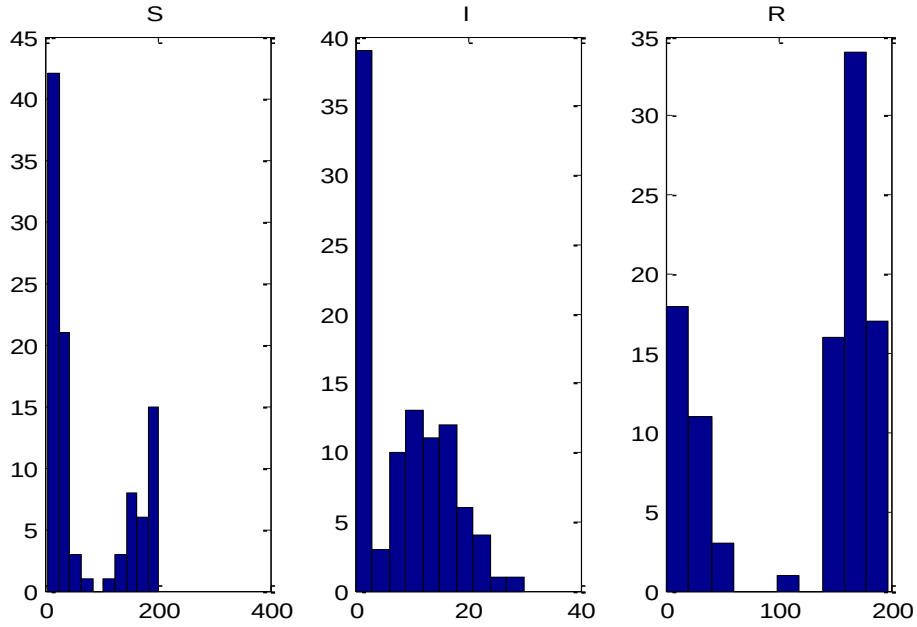



**Figure 20** – Histograms for 100 simulations with $R_0 = 10, \alpha_1 = 26, \alpha_2 = 2$, when $\mu = 1$, $\beta = 260, \gamma = 25, f = 1$; $S(0) = 20, I(0) = 13, R(0) = 167, N_0 = 200$ and $T = 2$.

The sample in **Figure 20** is still constructed by 100 simulations. The histograms are bimodal because there are more than 30 times the processes end earlier, which definitely makes the normal assumption unjustified and thus the QSD approximates the sample means and standard deviations,

$$\bar{S}(2) = 63.63, \bar{I}(2) = 8.91, \bar{R}(2) = 126.83,$$

$$S_s(2) = 69.19, I_s(2) = 7.20, R_s(2) = 64.57$$

incorrectly and meaninglessly.

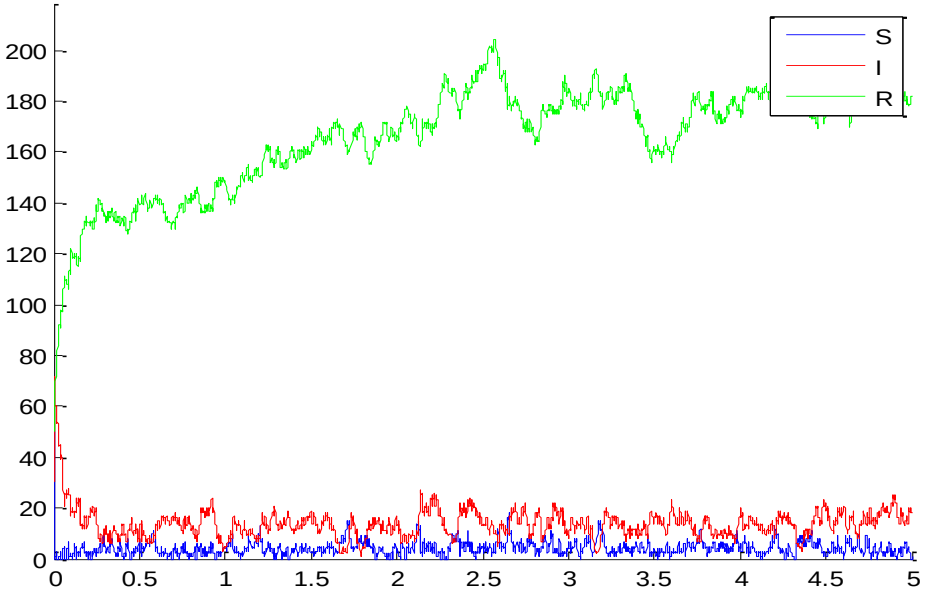



**Figure 21** – Simulation result with $R_0 = 50, \alpha_1 = 26, \alpha_2 = 2$, when $\mu = 1$, $\beta = 1300, \gamma = 25, f = 1$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 20$.

As $\alpha_1 = \gamma + 1 \to \infty$, the recovery parameter $\gamma$ is large under this condition, i.e. it takes relatively short time for an infective host to recover. In order to delay the extinction time of the disease without a small $\gamma$, $R_0$ have to be increased. Larger values of $R_0 = 15, 20, 25, 30, 35, 40, 45, 50$, etc. have the tendency to make the disease more endemic with overwhelmingly many recovered hosts. All the processes look similar as in **Figure 21**, but occasionally, the disease dies out early (as there are very large $S(2)$ which correspond to $I(2) = 0$ in **Figure 22**).

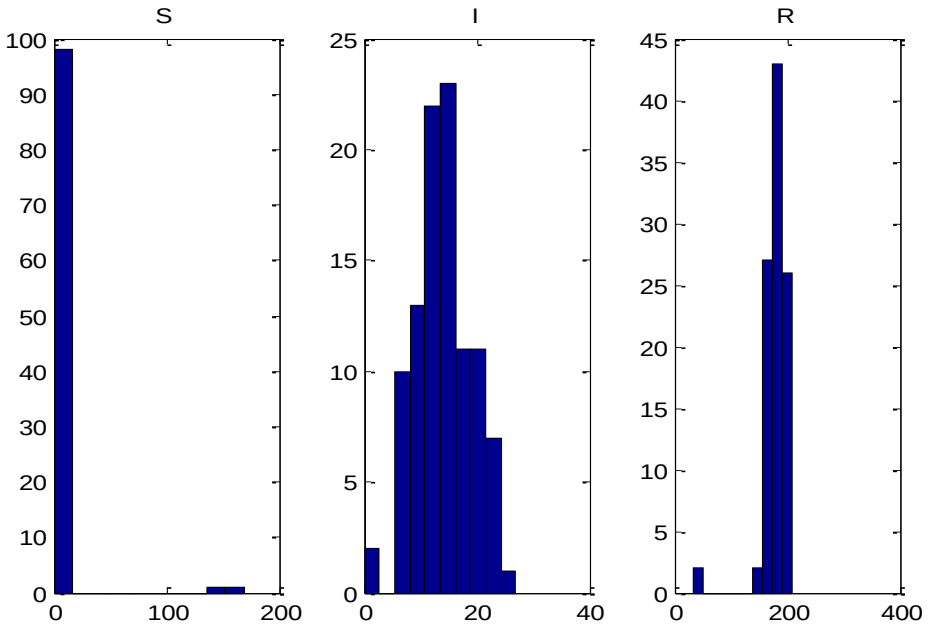



**Figure 22** – Histograms for 100 simulations with $R_0 = 50, \alpha_1 = 26, \alpha_2 = 2$, when $\mu = 1$, $\beta = 1300, \gamma = 25, f = 1$; $S(0) = 4, I(0) = 15, R(0) = 181, N_0 = 200$ and $T = 2$.

The QSD approximation:

$$\mu_S \approx 4.00, \sigma_S \approx 0.08,$$

$$\mu_I \approx 15.08, \sigma_I \approx 1.98,$$

$$\mu_R \approx 196, \sigma_R \approx 0.08$$

are quite compatible with **Figure 21**, though they disagree on the value of $R$. However, there are 2 simulations with $I(2) = 0$, due to which the values of $S(2), R(2)$ are far away from the majority of the data, shifting the means away from the modes and increasing the standard deviations:

$$\bar{S}(2) = 7.90, \bar{I}(2) = 14.06, \bar{R}(2) = 179.00,$$

$$S_s(2) = 22.02, I_s(2) = 4.93, R_s(2) = 23.87.$$

The simulation results are obviously different from the approximations, but since the normal assumption is not satisfied, this comparison is meaningless.

However, if the 2 simulations are treated as outliers, the approximation may provide reasonable results. The corrected means and standard deviations can be obtained:

$$\bar{S}(2) = 4.80, \bar{I}(2) = 14.35, \bar{R}(2) = 181.87,$$

$$S_s(2) = 2.54, I_s(2) = 4.55, R_s(2) = 13.20,$$

which is closer to the approximations. With proper statistical tests, I have

$$\mathrm{X_S} = 101711.73, \mathrm{X_I} = 511.79, \mathrm{X_R} = 2748105.10,$$

$$\mathrm{Z_S} = 100.44, \mathrm{Z_I} = -3.65, \mathrm{Z_R} = -1787.32.$$

Hence, the QSD approximation is still terrible, as it dramatically underestimates the standard deviations.

Nonetheless, it is interesting to notice the QSD approximation under the previous condition for the common cold:

$$\mu_S \approx 4.00, \sigma_S \approx 2.28,$$

$$\mu_I \approx 14.00, \sigma_I \approx 4.23,$$

$$\mu_R \approx 182.00, \sigma_R \approx 3.72$$

is better, which gives

$$X_S = 120.39, X_I = 112.19, X_R = 1222.97,$$

$$Z_S = 3.47, Z_I = 0.82, Z_R = -0.35.$$

Although the approximation estimates the value of $S$ unfairly, the results are considerably better. This phenomenon should suggest that $\alpha_1 = 26, \alpha_2 = 2$ is not compatible with the theoretical condition $\alpha_1 \to \infty, \alpha_2 = O(1)$, but for larger $\alpha_1$, the disease will die out even more quickly.

*4.3.4 Endemic chickenpox*

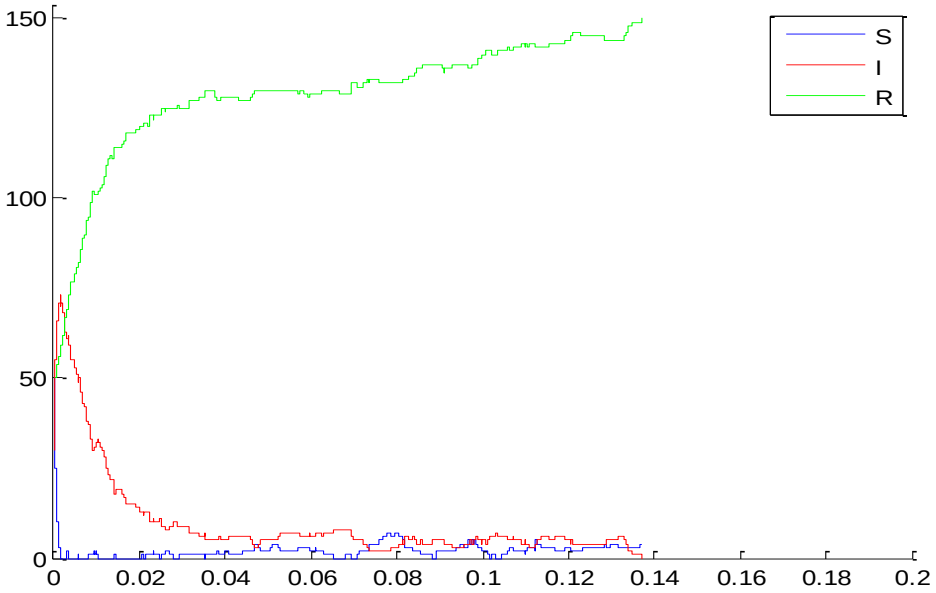



**Figure 23** – Simulation result with $R_0 = 50, \alpha_1 = 100, \alpha_2 = 2$, when $\mu = 1$, $\beta = 5000, \gamma = 99, f = 1$; $S(0) = 50, I(0) = 30, R(0) = 20, N_0 = 200$ and $T = 20$.

In the case shown by **Figure 23** where $\alpha_1 = 100$, the disease frequently dies out before $T = 0.2$. To delay the extinction, I increase $R_0$ to 500, even 5000, but it makes slightly differences. Since

$$\lim_{R_0 \to \infty} \hat{I} = \lim_{R_0 \to \infty} \frac{1+f}{1+f+\gamma} N_0 = \frac{400}{101} \approx 3.96,$$

$$\lim_{R_0\to\infty}\mu_I \approx \lim_{R_0\to\infty}\frac{\alpha_2}{\alpha_1}N_0 \approx 4,$$

the approximated number of infective hosts is close to zero, the absorbing state, which makes the process very likely to terminates. Hence, in order to make the approximations work, a relative large $\hat{I}$ or $\mu_I$ is necessary; the only parameter left for variation is $N_0$. In other words, we expect to have fair approximations by enlarging $N_0$, which also agrees with the requirement for a sufficiently large population.

Though the disease still dies quickly with $N_0$ doubled or tripled for $R_0 = 50$, the extinction time tends to be later and for $N_0 = 2000$, the disease appears to be endemic (see **Figure 24**). The QSD approximation in this case is:

$$\mu_S \approx 40.00, \sigma_S \approx 0.13,$$

$$\mu_I \approx 39.20, \sigma_I \approx 6.26,$$

$$\mu_R \approx 1960.00, \sigma_R \approx 0.13,$$

which suggests the number of recovered hosts might have not approached the equilibrium.

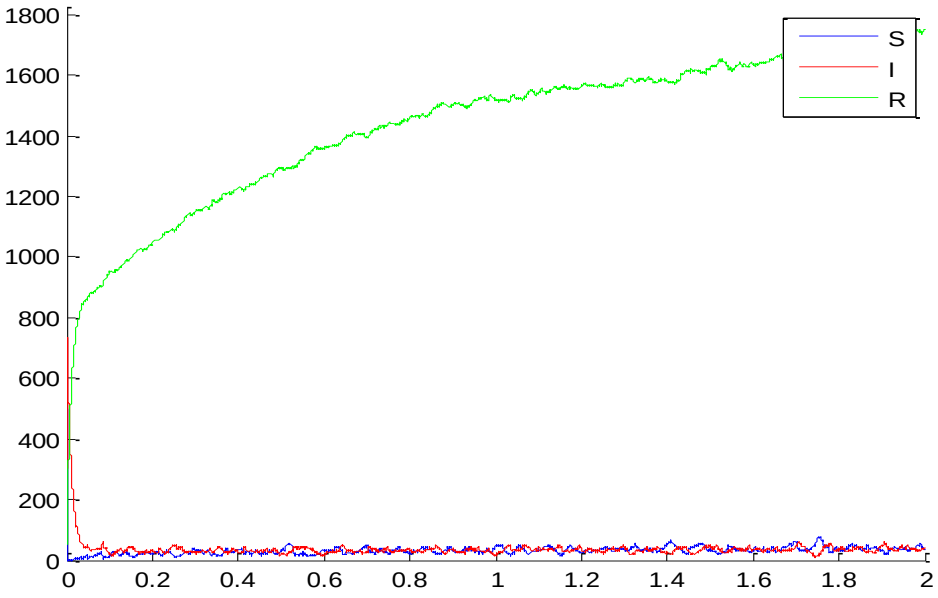



**Figure 24** –Simulation result with $R_0 = 50, \alpha_1 = 100, \alpha_2 = 2$, when $\mu = 1$, $\beta = 5000, \gamma = 99, f = 1$; $S(0) = 500, I(0) = 300, R(0) = 200, N_0 = 2000$ and $T = 2$.

A sample of 100 simulations gives

$$\bar{S}(2) = 42.07, \bar{I}(2) = 37.38, \bar{R}(2) = 1917.60,$$

$$S_s(2) = 9.94, I_s(2) = 10.04, R_s(2) = 44.26,$$

with **Figure 25**. The normal approximation seems to be justified but the QSD seems to underestimate

the standard deviations again. Indeed,

$$X_S = 59942.26, X_I = 249.64, X_R = 1187662.12,$$

$$Z_S = 162.00, Z_I = -2.88, Z_R = -3318.32.$$

suggests the QSD is not valid in this case.

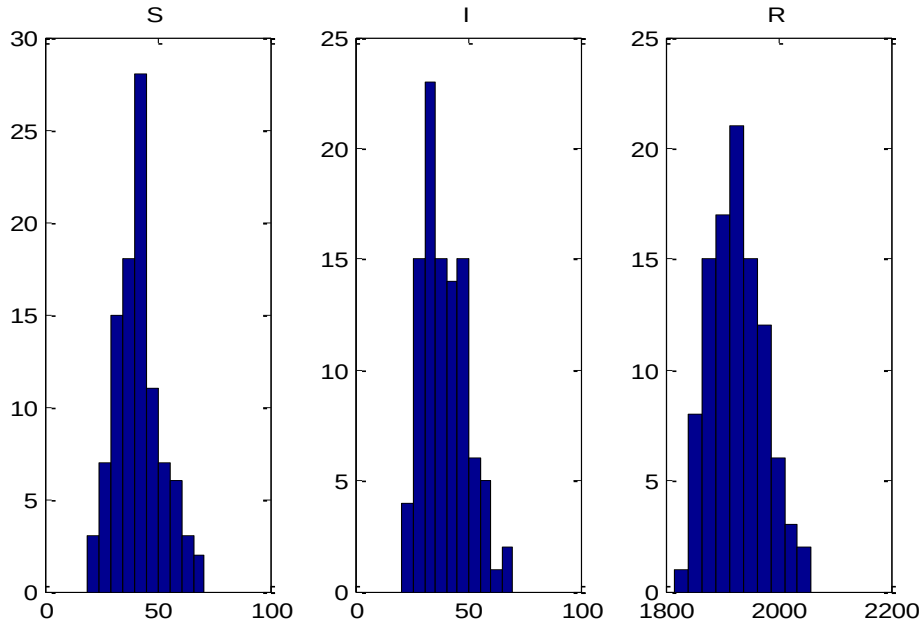



**Figure 25** – Histograms for 100 simulations with $R_0 = 50, \alpha_1 = 100, \alpha_2 = 2$, when $\mu = 1$, $\beta = 5000, \gamma = 99, f = 1$; $S(0) = 40, I(0) = 39, R(0) = 1921, N_0 = 2000$ and $T = 2$.

Notice again, the QSD approximates for the common cold:

$$\mu_S \approx 40.00, \sigma_S \approx 9.03,$$

$$\mu_I \approx 38.43, \sigma_I \approx 8.81,$$

$$\mu_R \approx 1921.57, \sigma_R \approx 6.20,$$

which gives

$$X_S = 117.56, X_I = 126.08, X_R = 4948.63,$$

$$Z_S = 2.27, Z_I = -1.18, Z_R = -6.33.$$

Although the approximation for the mean and standard deviation for the number of recovered hosts are not good, it is considerably much better than the QSD approximation for chickenpox.

Therefore, Nåsell's QSD approximation for diseases like chickenpox works unfairly, but the QSD approximation for disease like common cold works functionally in both the cases.

## *5. Conclusion*

This paper has worked on the epidemic process with the SIRS stochastic model with demography. There are debates over the transmission term, density-dependent transmission (DT) or frequency-dependent transmission (FT), but it seems very difficult to compare them analytically. Hence, computer simulations are conducted and the results agree with heuristic reasoning, which further suggests that the DT model requires a larger population than the FT model for its equivalent deterministic model to be a fair approximation.

As it is not realistic to derive any analytic solutions, the long-term analysis of equilibrium may help us understand the process. The deterministic equilibrium is trivial but the stochastic process has no explicit stationary distribution. Instead, the quasi-stationary distribution (QSD) is defined to be a reasonable alternative and Nåsell (2002) approximates it under two particular conditions.

The results of the computer simulations are then compared with these approximations. The QSD approximation for disease like chickenpox is invalid, but that for diseases like the common cold is fairly good, which, surprisingly, also works well for the first case. Additionally, the approximations are better with larger reproductive ratios.

In conclusion, Nåsell's QSD approximation for the common cold is an effective, efficient and robust tool in studying the FT model. It will be interesting to find similar approximations for the DT model and, by comparing them, a more general understanding of the stochastic epidemic could be obtained, which might make us more advantageous in reacting to or even controlling the disease.

## *Bibliography*

Andersson, H. and Britton, T. (2000). *Stochastic epidemic models and their statistical analysis* Springer-Verlag New York, Inc.

Bailey, N. T. J. (1957). *The mathematical theory of epidemics* Charles Griffin & Company Limited

Bartlett, M. S. (1949). Some evolutionary stochastic process. *J. Roy. Statist. Soc. B* **11**, 211-229.

Begon, M., Bennett, M., Bowers, R. G., French, N. P., Hazel, S. M., Turner, J. (2002). A clarification of transmission terms in host-microparasite models: numbers, densities and areas. *Epidemiol. Infect* **129**, 147-153.

Bernoulli, Daniel (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole, et des avantages de l'inoculation pour la prévenir. *Mém. Math. Phys. Acad. Roy. Sci., Paris*, 1-45. In *Histoire de l'Académie Royale des Sciences*, (1766).

Daley, D. J., Gani, J. (1999). *Epidemic Modelling: An Introduction* Cambridge University Press

Dawkins, C. R. (2006). *The Selfish Gene – 30th anniversary edition* Oxford University Press

Heffernan, J. M., Smith, R. J., Wahl, L.M. (2005). Perspectives on the basic reproductive ratio. *J. R. Soc. Interface* **2**, 281–293.

Kermack, W. O. and McKerdrick, A. G. (1927). A contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. Lond. A* **115**, 700-721

Kermack, W. O. and McKerdrick, A. G. (1932). Contributions to the mathematical theory of epidemics. II. the problem of endemicity. *Proc. Roy. Soc. Lond. A* **138**, 55-58

Kermack, W. O. and McKerdrick, A. G. (1933). Contributions to the mathematical theory of epidemics. III. further studies of the problem of endemicity. *Proc. Roy. Soc. Lond. A* **141**, 94-122

Li, M. Y., Muldowney, J. S. (1995). Global stability for the SEIR model in epidemiology. *Math. Biosci.* **125**, 155

Liu, W.-M., Hethcote, H. W., Levin, S. A. (1987). Dynamical behaviour of epidemiological models with nonlinear incidence rates. *J. Math. Biol.* **25**, 359-380

McKerdrick, A. G. (1926). Applications of mathematics to medical problems. *Proc. Edinburgh Math. Soc.* **14**, 98-130

Mollison, D. (1995). *Epidemic Models: Their Structure and Relation to Data* Cambridge University Press

Nåsell, I. (2002). Stochastic models of some endemic infections. *Mathematical Biosciences* **179**, 1–19.

Norris, J. R. (1997). *Markov Chains* Cambridge University Press

Ross, S. M. (2010). *Introduction to Probability Models – 10th ed.* Elsevier Inc.

Verhulst, P.-F. (1838). Notice sur la loi que la population poursuit dans son accroissement. *Corr. Math. Phys. (Bruxelles)* (publ. par A. Quételet) **10**, 113–121.

# *Appendix*

### SIRS_parameter.m

```
% values of parameters
clear,clc,clf
global mu beta gamma f N0
T=5,N0=2000,S0=500,I0=300,R0=200,N=S0+I0+R0,
%1:by specifying initial parameters
%mu=1,beta=100,gamma=25,f=100,
%R0=beta/(mu+gamma),alpha1=gamma/mu+1,alpha2=f/mu+1,
%2:by specifying reparameterisation
R0=50,alpha1=100,alpha2=2,
mu=1,beta=R0*alpha1*mu,gamma=(alpha1-1)*mu,f=(alpha2-1)*mu,
```

Remark: cf. *Section 4.3.1*; the values of the variables should change according to different cases.

### SIRS_equilibria.m

```
% equilibria of process
R02=R0^2;R03=R0^3;
theta=alpha2/alpha1;theta2=theta^2;theta3=theta^3;
%0:deterministic equilibrium
SIRdetmean=N0*[1/R0,(mu+f)/(mu+f+gamma)*(1-1/R0),gamma/(mu+f+gamma)*(1-1/R0
)]'
%1:QSD - nasell's 1st approximation
muS1=N0/R0;
muI1=N0*theta*(R0-1)/R0;
muR1=muI1/theta;
sigmaS1=N0*theta/R02;
sigmaI1=N0*(R0-1)/R02;
sigmaR1=sigmaS1;
SIRqsd1musigma=[muS1,sqrt(sigmaS1);muI1,sqrt(sigmaI1);muR1,sqrt(sigmaR1)]
%2:QSD - nasell's 2nd approximation
muS2=N0/R0;
muI2=N0*theta/(theta+1)*(R0-1)/R0;
muR2=muI2/theta;
sigmacommon=N0/R0^2/(theta+R0);
sigmaS2=sigmacommon/theta*((R0+1)*theta2+(R02+2*R0)*theta+R0);
sigmaI2=sigmacommon/(theta+1)^2*((R02-R0+1)*theta3+(R03+R0)*theta2+(R03+R02
-R0)*theta+R02-R0);
sigmaR2=sigmacommon/(theta+1)^2*((R02-R0)*theta3+(R03-R0+1)*theta2+(R03-2*R
02+3*R0)*theta+R0);
SIRqsd2musigma=[muS2,sqrt(sigmaS2);muI2,sqrt(sigmaI2);muR2,sqrt(sigmaR2)]
```

Remark: cf. *Section 3.3.2 & 3.4.3*; this m-file works after executing *SIRS_parameter.m*. successfully.

**SIRS_detcurve.m**

```
% curves of deterministic models
hold on
%DT
[ts,sir]=ode45(@detdt,[0,T],[S0;I0;R0]);
plot(ts,sir(:,1),'Color',[0 0 .4])
plot(ts,sir(:,2),'Color',[.6 0 0])
plot(ts,sir(:,3),'Color',[0 .5 0])
%FT
[ts,sir]=ode45(@detft,[0,T],[S0;I0;R0]);
plot(ts,sir(:,1),'Color',[0 0 1])
plot(ts,sir(:,2),'Color',[1 0 0])
plot(ts,sir(:,3),'Color',[0 1 0])
```

**detdt.m**

```
function dE=detdt(t,E)
global mu beta gamma f N0
dE=zeros(3,1);
dE(1)=mu*(N0-E(1))-beta*E(1)*E(2)/N0+f*E(3);
dE(2)=beta*E(1)*E(2)/N0-(mu+gamma)*E(2);
dE(3)=gamma*E(2)-(mu+f)*E(3);
end
```

**detft.m**

```
function dE=detft(t,E)
global mu beta gamma f N0
dE=zeros(3,1);
dE(1)=mu*(N0-E(1))-beta*E(1)*E(2)/(E(1)+E(2)+E(3))+f*E(3);
dE(2)=beta*E(1)*E(2)/(E(1)+E(2)+E(3))-(mu+gamma)*E(2);
dE(3)=gamma*E(2)-(mu+f)*E(3);
end
```

Remark: cf. *Section 3.3.1 & 4.1*; these m-files works after executing *SIRS_parameter.m*. successfully.

**SIRS_stochasticDT.m**

```
% simulation of stochastic DT model
S=S0;I=I0;R=R0;
Ts=zeros(1,500);
Ss=Ts;Ss(1)=S;
Is=Ts;Is(1)=I;
```

```
Rs=Ts;Rs(1)=R;
Ns=Ts;Ns(1)=N;
t=0;i=1;
while Ts(i)<=T
    Nbthr=mu*N0;
    Sdthr=mu*S;
    Idthr=mu*I;
    Rdthr=mu*R;
    SItrr=beta*S*I/N0;
    IRtrr=gamma*I;
    RStrr=f*R;
    total=Nbthr+Sdthr+Idthr+Rdthr+SItrr+IRtrr+RStrr;
    t=random('exp',1/total);
    Ts(i+1)=Ts(i)+t;
    p1=Nbthr/total;
    p2=Sdthr/total+p1;
    p3=Idthr/total+p2;
    p4=Rdthr/total+p3;
    p5=SItrr/total+p4;
    p6=IRtrr/total+p5;
    pr=rand;
    if pr<p1
        S=S+1;
    elseif pr<p2
        S=S-1;
    elseif pr<p3
        I=I-1;
    elseif pr<p4
        R=R-1;
    elseif pr<p5
        S=S-1;
        I=I+1;
    elseif pr<p6
        I=I-1;
        R=R+1;
    else
        S=S+1;
        R=R-1;
    end
    Ss(i+1)=S;
    Is(i+1)=I;
    Rs(i+1)=R;
    N=S+I+R;
    Ns(i+1)=N;
    i=i+1;
    if I==0
```

```
        break;
    end
end
Ts=Ts(1:i);
Ss=Ss(1:i);
Is=Is(1:i);
Rs=Rs(1:i);
hold on
axis([0 T 0 max(Ns)])
stairs(Ts,Ss,'Color',[0 0 .4]),stairs(Ts,Is,'Color',[.6 0
0]),stairs(Ts,Rs,'Color',[0 .5 0])
```

### SIRS_stochasticFT.m

```
% simulation of stochastic FT model
S=S0;I=I0;R=R0;
Ts=zeros(1,500);
Ss=Ts;Ss(1)=S;
Is=Ts;Is(1)=I;
Rs=Ts;Rs(1)=R;
Ns=Ts;Ns(1)=N;
t=0;i=1;
while Ts(i)<=T
    Nbthr=mu*N0;
    Sdthr=mu*S;
    Idthr=mu*I;
    Rdthr=mu*R;
    SItrr=beta*S*I/N;
    IRtrr=gamma*I;
    RStrr=f*R;
    total=Nbthr+Sdthr+Idthr+Rdthr+SItrr+IRtrr+RStrr;
    t=random('exp',1/total);
    Ts(i+1)=Ts(i)+t;
    p1=Nbthr/total;
    p2=Sdthr/total+p1;
    p3=Idthr/total+p2;
    p4=Rdthr/total+p3;
    p5=SItrr/total+p4;
    p6=IRtrr/total+p5;
    pr=rand;
    if pr<p1
        S=S+1;
    elseif pr<p2
        S=S-1;
    elseif pr<p3
        I=I-1;
    elseif pr<p4
```

```
        R=R-1;
    elseif pr<p5
        S=S-1;
        I=I+1;
    elseif pr<p6
        I=I-1;
        R=R+1;
    else
        S=S+1;
        R=R-1;
    end
    Ss(i+1)=S;
    Is(i+1)=I;
    Rs(i+1)=R;
    N=S+I+R;
    Ns(i+1)=N;
    i=i+1;
    if I==0
        break;
    end
end
Ts=Ts(1:i);
Ss=Ss(1:i);
Is=Is(1:i);
Rs=Rs(1:i);
hold on
axis([0 T 0 max(Ns)])
stairs(Ts,Ss,'Color',[0 0 1]),stairs(Ts,Is,'Color',[1 0
0]),stairs(Ts,Rs,'Color',[0 1 0])
legend('S','I','R')
```

Remark: cf. *Section 3.4.1, 3.4.2 & 4.1*; these two m-files works after executing *SIRS_parameter.m*. successfully; the only difference between them is inside the while-loop: `SItrr=beta*S*I/N` and `SItrr=beta*S*I/N0`; hence, they are for the two different transmission terms, respectively.

### SIRS_histogram.m

```
% 100 simulations
S0=round(SIRdetmean(1));I0=round(SIRdetmean(2));R0=round(SIRdetmean(3));
J=100;
for j=1:J
S=S0;I=I0;R=R0;N=S+I+R;
global mu beta gamma f N0
T=2;
Ts=0;i=1;
while Ts<=T
```

```
    Nbthr=mu*N0;
    Sdthr=mu*S;
    Idthr=mu*I;
    Rdthr=mu*R;
    SItrr=beta*S*I/N;
    IRtrr=gamma*I;
    RStrr=f*R;
    total=Nbthr+Sdthr+Idthr+Rdthr+SItrr+IRtrr+RStrr;
    t=random('exp',1/total);
    Ts=Ts+t;
    p1=Nbthr/total;
    p2=Sdthr/total+p1;
    p3=Idthr/total+p2;
    p4=Rdthr/total+p3;
    p5=SItrr/total+p4;
    p6=IRtrr/total+p5;
    pr=rand;
    if pr<p1
        S=S+1;
    elseif pr<p2
        S=S-1;
    elseif pr<p3
        I=I-1;
    elseif pr<p4
        R=R-1;
    elseif pr<p5
        S=S-1;
        I=I+1;
    elseif pr<p6
        I=I-1;
        R=R+1;
    else
        S=S+1;
        R=R-1;
    end
    N=S+I+R;
    i=i+1;
end
Sh(j)=S;
Ih(j)=I;
Rh(j)=R;
end
Inonzero=find(Ih~=0);n=length(Inonzero)
Sh=Sh(Inonzero);Ih=Ih(Inonzero);Rh=Rh(Inonzero);
SIR_mean_std=[mean(Sh),std(Sh);mean(Ih),std(Ih);mean(Rh),std(Rh)]
subplot(1,3,1),hist(Sh),title('S')
```

```
subplot(1,3,2),hist(Ih),title('I')
subplot(1,3,3),hist(Rh),title('R')
```

Remark: cf. *Section 4.2.1*; this m-file works after executing *SIRS_parameter.m & SIRS_equilibria.m* successfully.

### SIRS_statistests.m

```
% statistical tests for theoretical approximations
histmeanSIR=SIR_mean_std(:,1);
histstdvSIR=SIR_mean_std(:,2);
TSIR=(histmeanSIR-SIRdetmean)./histstdvSIR*sqrt(n)
%quasiSIR=SIRqsd1musigma;
quasiSIR=SIRqsd2musigma;
quasimeanSIR=quasiSIR(:,1);
quasistdvSIR=quasiSIR(:,2);
XSIR=(n-1)*(histstdvSIR./quasistdvSIR).^2
ZSIR=(histmeanSIR-quasimeanSIR)./quasistdvSIR*sqrt(n)
```

Remark: cf. *Section 4.2.2*; this m-file works after executing *SIRS_parameter.m, SIRS_equilibria.m & SIRS_histogram* successfully.