Modeling Statistical phenomena
Assignment 5

Ali Setareh Kokab (95100491)

Department of Physics, Sharif University of Technology

June 11, 2021

Contents

1 Ex. 1

1.1

The three distinguishing features of stochastic models are:

1. **Variability between simulations**: In stochastic models, each simulation causes a different outcome. This implies that although general statistical properties (such as the mean and the variance) may be accurately predicted, it is generally impossible to predetermine the precise disease prevalence at any given point in the future. However, in the deterministic models, the outcomes of the simulations are always the same (with the given set of parameters).

2. **Variances and covariances**: due to the existence of chance in the stochastic models, there is always a variance in the prevalence of the disease and the number of susceptibles. Moreover, the covariance between the number of infectious and susceptible individuals is negative, which causes the mean and variance of these two groups to be different from the deterministic model.

3. **Increased transients**: The underlying deterministic dynamics of the system always force the stochastic perturbations away from the equilibrium to get back to the equilibrium point. So the dynamics of stochastic models can be thought of as a random perturbation away from, and transient-like return toward, the deterministic attractor.

Using stochastic models will be most important whenever the number of infectious individuals is relatively small, which can be when the **population size is small**, when an **infectious disease has just invaded**, **when control measures are successfully applied**, or **during the trough phase of an epidemic cycle**.
One of the places that stochastic modeling is beneficial is in economics. For instance, to predict the price of a share, we should consider many factors from the behaviour of the shareholders to the Political situation of that country. So as it is obvious, most of these details are not available. If so, the relation between these factors is such complex that it is almost impossible to model them in a deterministic way. Moreover, we do not need to know the precise price of a share to decide whether or not to buy it, But it is sufficient to know the probability of that share in a given time horizon. So, instead of modeling the share price deterministically, we examine the general statistical features of that share. For example, in the minute time frame, we know the probability distribution of logarithm of the return of the price is a power law, and it shows scaling behaviors. However, in longer time frames, this distribution becomes gaussian. So by knowing the statistical features of this phenomenon, we could predict the chance that that share has a particular price. Even though we can not say much about the precise price of that share.

1.2

You see the system of equation in eq 1. X, Y, and Z represent the susceptibles, infected and recovered group. ξ_1 to ξ_6 are six distinct random numbers with Gaussian probability distribution.

$$\begin{aligned}
\frac{dX}{dt} &= [\nu N + \sqrt{\nu N}\xi_1] - [\beta XY/N + \sqrt{\beta XY/N}\xi_2] - [\mu X + \sqrt{\mu X}\xi_3] \\
\frac{dY}{dt} &= [\beta XY/N + \sqrt{\beta XY/N}\xi_2] - [\gamma Y + \sqrt{\gamma Y}\xi_4] - [\mu Y + \sqrt{\mu Y}\xi_5] \\
\frac{dZ}{dt} &= [\gamma Y + \sqrt{\gamma Y}\xi_4] - [\mu Z + \sqrt{\mu Z}\xi_6]
\end{aligned} \tag{1}$$

1.3

Now we solve the eq 1 for $\beta = 1.3, \gamma = \frac{1}{8}, \frac{1}{\mu} = \frac{1}{\nu} = 60(peryear)$. You can see the result in fig 1 and 2 for two different values of N_0, the total population. As you can see, for large population system acts as a deterministic model, and for a smaller population, we can see the stochasticity effects. We can see that the system reaches its equilibrium after a relatively long phase of oscillations for the larger population. We also observe that the system gets back to the deterministic equilibrium for a smaller population and larger experienced noise when it gets far from it. In other words, the underlying deterministic equations act like restorative forces and do not let the system get far from the equilibrium point.

$$\begin{aligned}
\frac{dX_H}{dt} &= \nu_H - rT_{HM}Y_M X_H - \mu_H X_H, \\
\frac{dY_H}{dt} &= rT_{HM}Y_M X_H - \mu_H Y_H - \gamma_H Y_H, \\
\frac{dX_M}{dt} &= \nu_M - rT_{MH}Y_H X_M - \mu_M X_M, \\
\frac{dY_M}{dt} &= rT_{MH}Y_H X_M - \mu_M Y_M,
\end{aligned} \tag{2}$$

1.4

In fig 3 you see the mean versus variance of infected group for different population size of SIR system with scaled noise. For scaled noise, and large population sizes (with more that 500 infected individuals on average), the variance scales linearly with the mean and therefore with the population size as well.Thus, although large populations behave more like the deterministic equations, the absolute level of variation is greater than in small populations. For very small populations, this linear relationship between the mean and variance is vanished because of the strong nonlinearities that operate when relatively large amplitude epidemics are triggered. As you can see in this plot, the fitted line in the

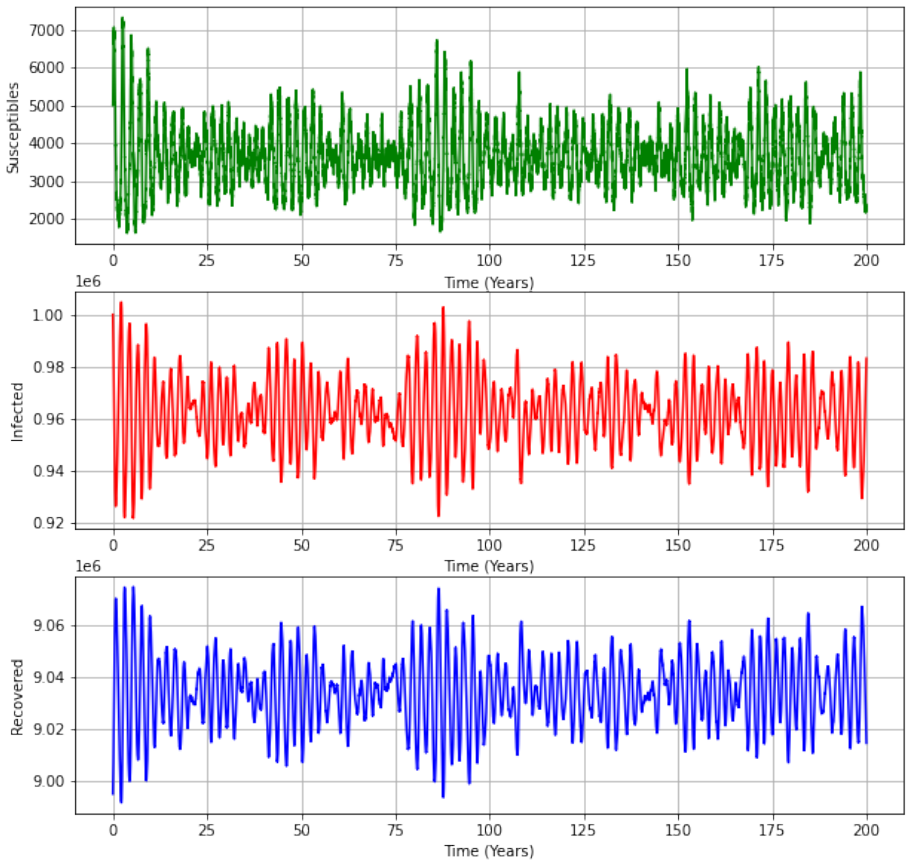

Figure 1: The dynamics of SIR epidemics with scaled noise. $\mu = \nu = \frac{1}{\mu} = 60 years$, $1/\gamma = 8$ days and $\beta = 1.3$ for N_0=10000000. initial values are X_0=1000000 ,Y_0=5000

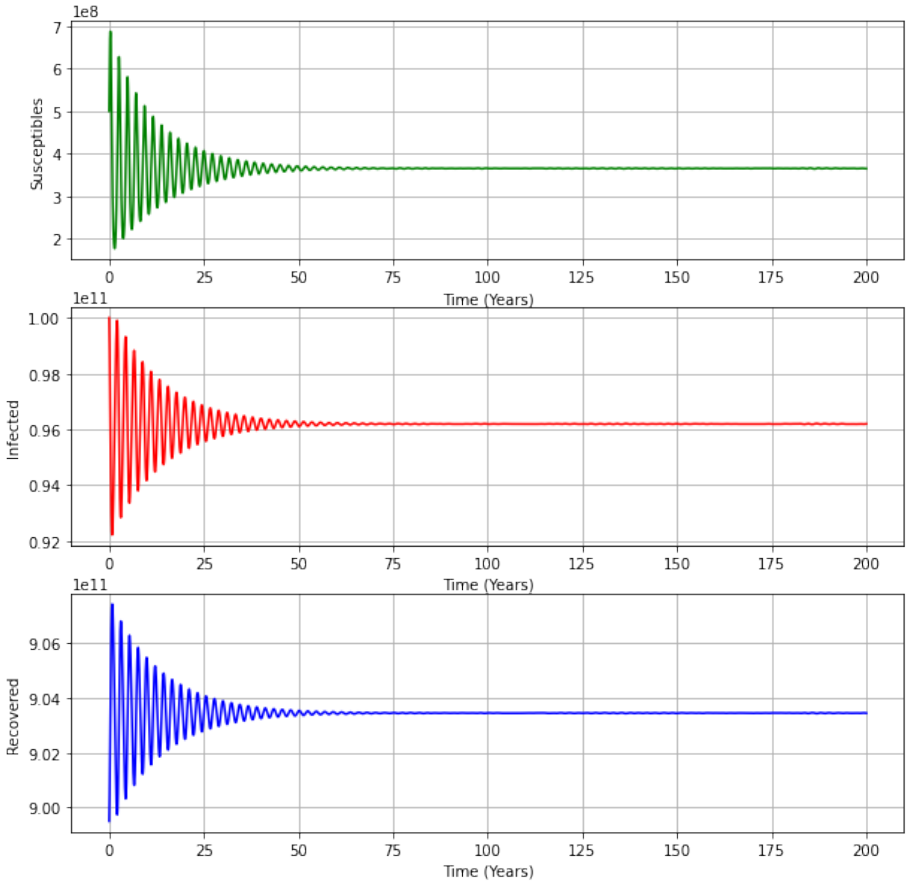

Figure 2: The dynamics of SIR epidemics with scaled noise. $\mu = \nu = \frac{1}{\mu} = 60 years$, $1/\gamma = 8$ days and $\beta = 1.3$ for N_0=1000000000000. initial values are X_0=100000000000 ,Y_0=500000000

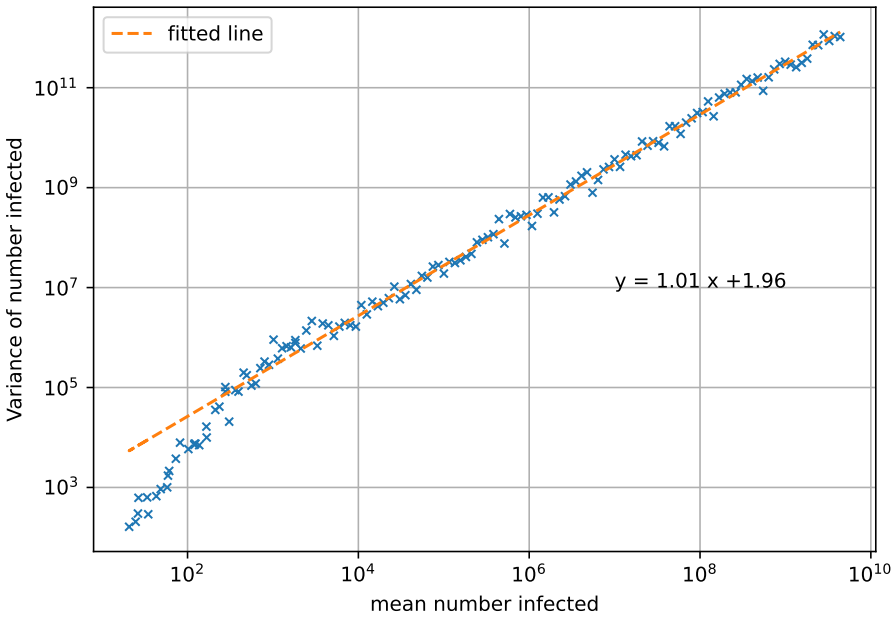

Figure 3: mean variance relationship for a range of population sizes from 100 to 10000 million. The dashed line represents var = 91 × mean, showing how the scaling operates at large population sizes.

log log scale has an intercept of 1.96, which determines that the variance and mean are related with var = 91 mean.

2 Ex. 2

In this exercise, we examine the SIR model with demographic stochasticity. We can simulate this behavior with event-driven methods.

2.1

Event-driven methods require explicit consideration of events. For the standard SIR model, we need to consider the six events that can occur, each causing the numbers in the relative classes to increase or decrease by one:

1. Births occur at rate μN. Result: $X \to X + 1$.
2. Transmission occurs at rate $\beta \frac{XY}{N}$ Result: $Y \to Y + 1$ and $X \to X - 1$.
3. Recovery occurs at rate γY. Result: $Z \to Z + 1$ and $Y \to Y - 1$.

4. Deaths of X occur at rate μX Result $X \to X - 1$.
5. Deaths of Y occur at rate μY Result $Y \to Y - 1$.
6. Deaths of Z occur at rate μZ Result $Z \to Z - 1$.

2.2

In fig 4 and fig 6 you see the simulation results of this model with Gillespie's direct algorithm for a population size of 100000 and 50, respectively. As you can see, in fig 6. We can see the different events more clearly. In fig 4 and 5 you see the simulation result for the same parameters. In comparison to the deterministic SIR model, we see that different simulations give rise to different outcomes. This implies that it is generally impossible to predetermine the precise disease prevalence at any given point in the future. Another thing we could say is that in some simulations, the disease can be prevalent successfully, but in others5, the disease dies. We could also see from this figure that there is no upper bound for different groups in contrast to the SIR deterministic model. This is because it is possible that for multiple time steps, a susceptible be born, but none of them gets the disease. We could also see that in fig 6 due to small population, the disease dies very soon.

3 Ex. 3

Let $p_{SIR}(t)$ be the probability of being in the state (S;I;R) at time t and N = S+I+R. twelve processes can occur that modify the proportion of simulations in this state:

- positive terms:
    1. A susceptible can be born with rate $\mu(N-1)$ in a simulation in the state (S-1,I,R) such that there are now S susceptibles.
    2. A susceptible can die with rate $\mu(N+1)$ in a simulation in the state (S+1,I,R) such that there are now S susceptibles.
    3. A simulation in state I-1 and S+1 can have a susceptible individual become infected at rate $\beta\frac{I-1}{N}(S+1)$, such that there are now I infected and S susceptibles.
    4. A simulation in the state (S,I+1,R-1) can have an infected individual recover at rate $\gamma(I+1)$, such that I infected and R and recovered remain.
    5. An infected can die with rate $\mu(I+1)$ in a simulation in the state (S,I+1,R) such that there are now I infected.
    6. A recovered can die with rate $\mu(R+1)$ in a simulation in the state (S,I,R+1) such that there are now R infected.
- negative terms:

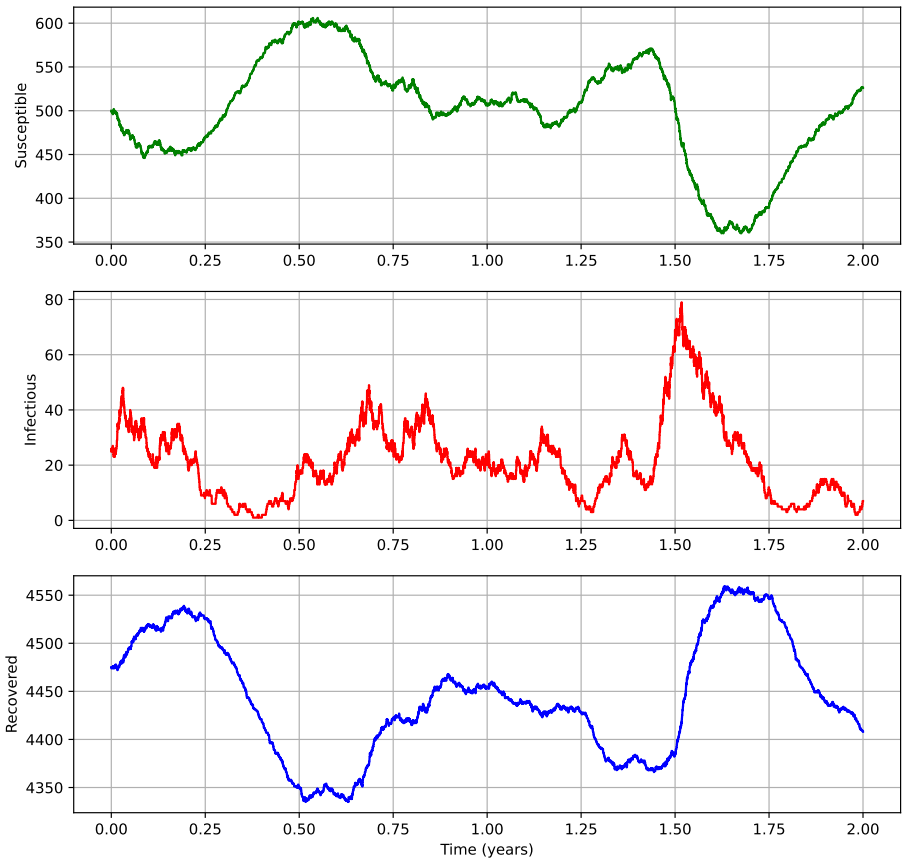

Figure 4: SIR model including births and deaths with full event-driven demographic stochasticity simulation with Gillespie's direct algorithm for population size of 100000. The parameters are: $\beta = 1.3$, $\gamma = \frac{1}{8}, \frac{1}{\mu} = 60$ (years) other parameters are in days. The initial values are $Y_0 = \text{ceil}(\mu \times N_0/\gamma) + 10$ $X_0 = floor(\gamma * N_0/\beta)$.

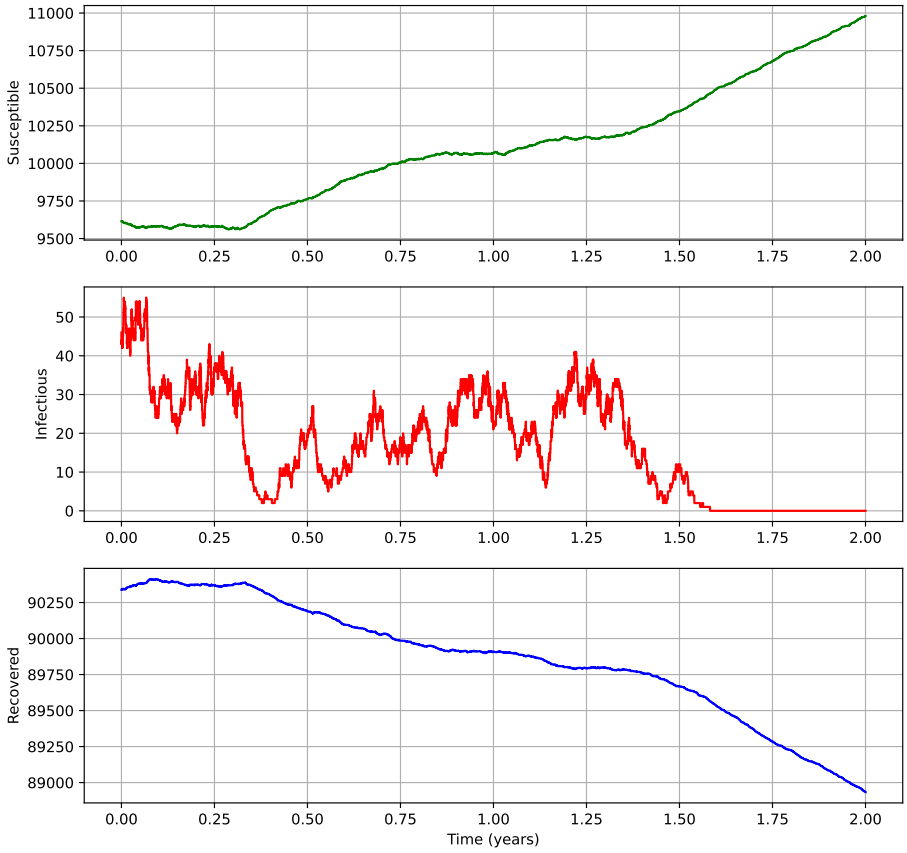

Figure 5: SIR model including births and deaths with full event-driven demographic stochasticity simulation with Gillespie's direct algorithm for population size of 100000 for the same parameters and initial values of fig 4. As you can see, due to the stochasticity of the model, different realizations have different outcomes.

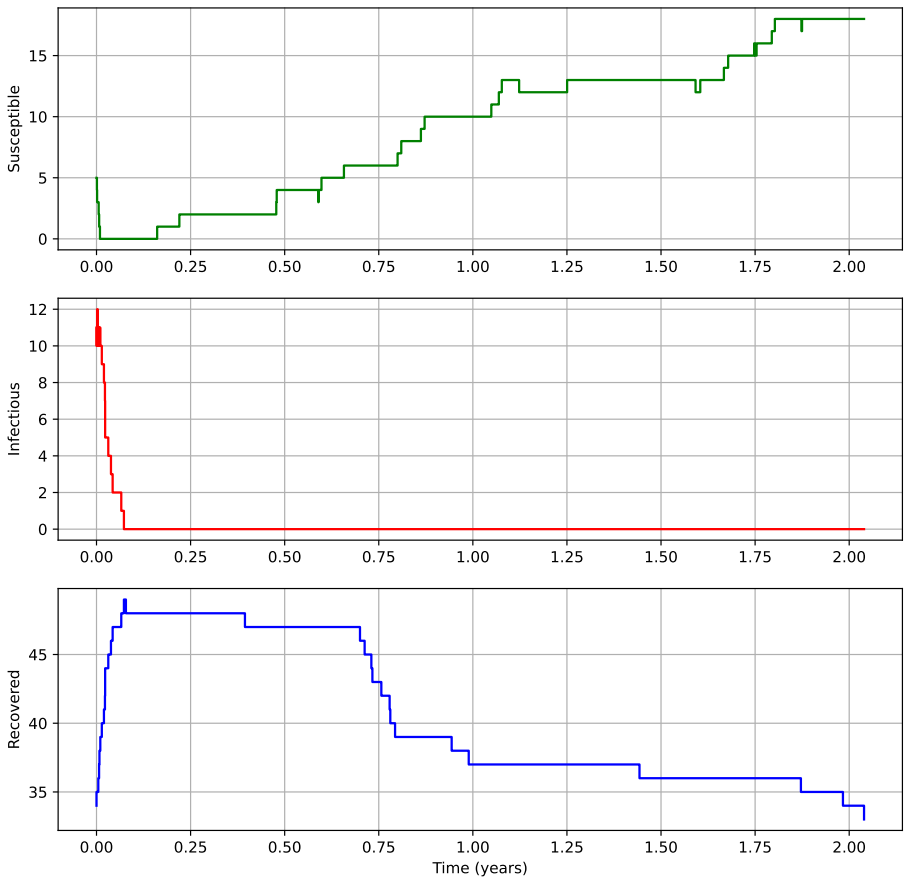

Figure 6: SIR model including births and deaths with full event-driven demographic stochasticity simulation with Gillespie's direct algorithm for population size of 50. The parameters are: $\beta = 1.3$, $\gamma = \frac{1}{8}$, $\frac{1}{\mu} = 60$ (years) other parameters are in days. The initial values are $Y_0 = \text{ceil}(\mu \times N_0/\gamma) + 10$ $X_0 = floor(\gamma * N_0/\beta)$. As you can see, in comparison with fig 6, in this plot the events are visible.

1. A susceptible can die with rate μS in a simulation in the state (S,I,R) such that there are now S-1 susceptibles.
2. A susceptible can be born with rate μN in a simulation in the state (S,I,R) such that there are now S+1 susceptibles.
3. A simulation in the state I and S can have a susceptible individual become infected at rate $\beta \frac{I}{N}(S)$, such that there are now I+1 infected and S-1 susceptibles.
4. A simulation in the state (S,I,R) can have an infected individual recover at rate $\gamma(I)$, such that $I-1$ infected and $R+1$ recovered remain.
5. An infected can die with rate $\mu(I)$ in a simulation in the state (S,I,R) such that there are now I-1 infected.
6. A recovered can die with rate $\mu(R)$ in a simulation in the state (S,I,R) such that there are now R-1 recovered.

Note that we assume that both death rate and birth rate are equal to μ. The following equation describes how this probability distribution evolves over time:

$$\begin{aligned}\frac{dp_{SIR}}{dt} = & p_{S-1,I,R}[\mu(N-1)] + p_{s+1,I,R}[\mu(S+1)] + p_{s+1,I-1,R}[\beta\frac{I-1}{N}(S+1)] \\ & + P_{S,I+1,R-1}[\gamma(I+1)] + P_{S,I+1,R}[\mu(I+1)] + p_{S,I,R+1}[\mu(R+1)] \\ & - p_{S,I,R}[\mu N + \mu S + \beta\frac{I}{N}S + \gamma I + \mu I + \mu R].\end{aligned} \tag{3}$$

As you can see the master equations for the SIR model has $\frac{1}{2}(N+1)(N+2)$ equations. The reason behind this is that for every value of $S \in \{0,...,N\}$, I can take all the values between 0 and N-S (N-S+1 values). So the number of equations are:

$$\sum_{i=0}^{N} N - i + 1 = \frac{1}{2}(N+1)(N+2)$$