

Incorporating stochasticity in simple models of disease spread

Jonathan Dushoff

As we have seen in earlier chapters, simple models of infectious diseases in homogeneous populations have produced many important insights. The most basic of these is the basic reproductive number R_0 , and its corollaries: the population (or ‘herd’) immunity threshold; and the qualitative shape of the relationship between R_0 , disease parameters and endemic levels of disease. Another key idea is the cyclic tendency of epidemic outbreaks, and the natural tendency of these oscillations to be damped.

Although such models have produced important insights, it is important to bear in mind that they are simple (albeit useful) approximations to a much more complex reality. Thus, it is also useful to consider what the key simplifications of these models are, and ways in which their predictions might change when various complications are added. Several such complications may be loosely grouped under the classification of “heterogeneities”: populations in the simple models we have learned about are fundamentally homogeneous.

Examples of heterogeneities not accounted for in simple models include “parametric heterogeneity”, “geographic heterogeneity” and “demographic heterogeneity.” Parametric heterogeneity refers to individual differences in transmission parameters for a particular disease; for example, differences among individuals in contact rate, susceptibility to infection, or tendency to transmit disease. Starting with the work of Hethcote and Yorke [11], parametric heterogeneity has been shown to pervasively affect the relationship between risk factors and disease – on the whole, tending to make disease levels less sensitive to increases and decreases in risk [8].

Geographic heterogeneity refers to the geometry of interaction networks between people; people do not mix homogeneously, but instead mixing is structured by spatial location, social networks, and other factors. The effects of geographic heterogeneity on disease spread is an active research area, which is further divided (loosely) into spatial models (see, e.g., [15]) and network models (e.g., [4]). This chapter will focus on demographic heterogeneity. A good source for further reading on this subject is [1].

This work was supported by the National Science and Engineering Research Council of Canada.

1. Modeling demographic heterogeneity

Demographic heterogeneity refers to the heterogeneity introduced simply by treating individuals as individuals. To see what this means, we first ask what happens in a classic, ODE model of disease, if we introduce a single infectious individual to a susceptible population. If $R_0 > 1$, after a short period of time, we deterministically find something like 1.1 infectious individuals, whereas if $R_0 < 1$, after a short period of time we have 0.9 infectious individuals.

In reality, of course, a single infectious individual does not increase smoothly to 1.1 infectious individuals. Instead, the system will take a discrete jump, either to 2 infectious individuals or to 0. We cannot reasonably expect to specify a model of disease transmission in enough detail that we can predict exactly when the next event of each type will occur. Instead, simple models that include discrete events are fundamentally stochastic; even in theory we don't know when the next event will occur, nor even what it will be.

Thus, if we treat individuals as individuals, we necessarily introduce “demographic” stochasticity, defined as stochasticity caused by the existence of individual people and discrete events. Stochasticity may also operate at a larger scale; we refer to events outside our model that affect more than one person at a time as “environmental” stochasticity. In a model of disease spread in human populations, sources of environmental stochasticity would include weather and economic change, for example.

This is illustrated in Figure 1. The deterministic trajectory starts from a single introduction and increases smoothly to around 2.5 infectious individuals. In contrast, the stochastic trajectories move in discrete, unpredictable jumps. Some of them randomly increase to as high as 8 or 9 infectious individuals, while in one case, the disease goes extinct before the end of the week.

2. Describing a demographic model

Interestingly, there is a natural way of translating a simple ODE model directly into an analogous discrete-individual model. This being so, it is useful to investigate how the behavior and predictions SIR models are changed by the inclusion of this realistic effect.

To create a discrete-events model, we describe our system in terms of the *probability rates* of events happening. If the rate of event E is $r_E(t)$, the probability of the event occurring in the time interval $(t, t + dt)$ is $r_E(t)dt$. If the system is *Markovian*, $r_E(t)$ depends only on the state of the system at time t . If the system is *non-Markovian* $r_E(t)$ can depend (for example) on how long the system has been in a particular state.

The Markovian assumption allows us to specify the state of our system with a much smaller amount of information, but it can have unwanted consequences. In particular, assuming that residence in a particular state is independent of history is equivalent to assuming that the residence time in that state is exponentially distributed. This assumption is often strikingly unrealistic – for example, in a human population with exponentially distributed lifetimes and a mean lifetime of 70 years, 5% of people will live to be over 200 years old! Thus, we must question the extent to which our results depend on this extremely convenient, but unrealistic, assumption.

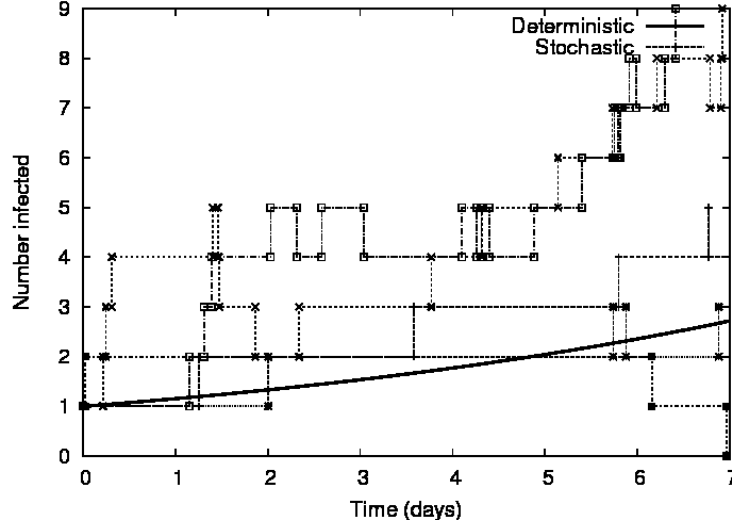


FIGURE 1. The deterministic trajectory from a simple SIR model (2.1) is compared with several trajectories from an analogous stochastic model (Table 1). Parameters are $R_0 = 2$, $\rho = 0.01$. The disease generation time is 7 days.

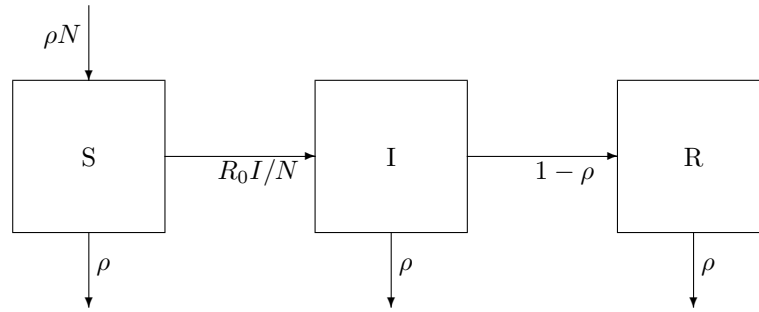


FIGURE 2. A diagrammatic model of the spread of an infectious disease. Individuals may be **S**usceptible, **I**nfectious, or **R**esistant. Susceptible individuals move to the infectious category at a rate proportional to the number of infectious individuals present. Births, deaths and recovery from disease into the Resistant class occur at constant rates.

Figure 2 shows a diagram of an infectious disease model. By counting time in units of infectious disease generations, and populations as fractions of the total population, we can reduce this simple, but by no means trivial, model to only two parameters, the basic reproductive number R_0 ; and ρ , which represents the ratio between the two fundamental time scales of the model: the average infectious period, and the average duration of immunity following infection.

$$(2.1) \quad \begin{aligned} \dot{x} &= \rho(1 - x) - R_0xy \\ \dot{y} &= R_0xy - y \end{aligned}$$

Event	transition	rate	Effect (S, I)
Infection	$S \rightarrow I$	$R_0 SI/N$	$(-1, 1)$
Recovery	$I \rightarrow R$	$(1 - \rho)I$	$(0, -1)$
Rebirth	$R \rightarrow S$	$\rho(N - S - I)$	$(1, 0)$
Rebirth	$I \rightarrow S$	ρI	$(1, -1)$

TABLE 1. We specify a demographic model by describing the events that can occur; the way that the rates of these events depend on the state of the system; and the effect of these events on the variables that describe the state of our system.

Here, x ; y refer to the proportion of the population infectious; susceptible. Since we assume a constant population, we are free to drop the equation for the proportion in the removed class. The dots refer to derivatives with respect to time.

Table 1 shows how we re-cast the model in Figure 2 as a demographic model, reflecting a population composed of discrete individuals. Each arrow represents an event that can occur stochastically, affecting a discrete individual, with a probabilistic rate that corresponds to the deterministic rate in (2.1). We will discuss below how this stochastic model of the disease system in Figure 2 can be either simulated, or analyzed. Note that the model given by Table 1 cannot be parameterized as simply as the one in (2.1); if individuals are discrete, the population size N will fundamentally affect the behavior of the model.

The demographic model is an exact analogue of the deterministic one, both conceptually, and in the technical sense that as the population size $N \rightarrow \infty$ in the demographic model, its behavior converges to that of the deterministic model. Although we can convert a demographic model to a deterministic model simply by allowing the population size to go to infinity, translating a deterministic model to a demographic one requires us to make choices about how to group the terms in our model into events. For example, the two transmission terms in our simple model – the contribution of $-R_0xy$ to \dot{x} , and the contribution of R_0xy to \dot{y} – clearly refer to the same event: an individual moves from the susceptible class to the infectious class. Less clear is the interpretation of birth and death terms. We could choose to interpret the model (2.1) as having independent birth and death events – and thus, stochastic fluctuations in population size – or we can choose (as we do in Table 1) to link these birth and death events, and keep the population constant.

Having described a model that takes demographic stochasticity into account, we can now begin to ask questions about the expected effects of demographic stochasticity on disease dynamics. In particular, we can now ask questions about the *probability* of establishment or persistence of diseases with $R_0 > 1$. We can also ask how we expect demographic stochasticity to affect temporal dynamics of diseases, and how much variability we expect to see due to demographic stochasticity.

3. The behavior of a stochastic model

We have experience in interpreting the behavior of deterministic models. Given parameters, and a starting point, we examine the predicted trajectory of the model. We also have techniques for combining information from different starting points into phase planes, and information using different parameter sets into bifurcation diagrams.

Simulating a realization

- Given a state of the system:
 - List possible events, and associated rates.
 - Randomly select the time and nature of the next event:
 - * Calculate the total rate r_T
 - * the next event will happen after an exponential waiting time with mean $1/r_T$.
 - * The probability of event E is r_E/r_T .
 - Change the system state appropriately.
- Repeat forever.
- Or until disease is extinct.
- Or until you are tired

TABLE 2. Steps in simulating a demographic model.

For stochastic models, we must re-evaluate what we mean by the “behavior” of a system. Given parameters, and a starting point, we do not know what the trajectory of the model will be. We can attempt to simulate a random sample of what the model might do. Such an example of a single trajectory is called a *realization*. We can attempt to characterize the universe of all possible such realizations; this is called the *ensemble*; we can imagine the ensemble as an infinity of realizations running simultaneously. Less ambitious than attempting to characterize the whole ensemble, we could ask about the probability distribution that describes what state we expect the population to be in at any given time; this is called the *ensemble distribution*.

We have a variety of techniques available to interpret a stochastic dynamical system. 1) Most straightforwardly, we can simulate one or many realizations. 2) We can simulate the entire ensemble distribution; this technique is restricted to systems with relatively small population size, since such simulations require that we keep track of one state variable for each possible state of the system. 3) We can solve the ensemble distribution dynamics exactly; this is possible only in special cases, but these solutions can be enlightening. 4) We can construct analytic *approximations* to the ensemble distribution.

3.1. Simulating individual realizations. Continuous-time models with discrete events can be simulated in a straightforward manner, outlined below Table 2. This method has been known for at least half a century, and is attributed to Kendall by [6]. It was explicated in detail by Gillespie [10], and is often referred to as the Gillespie method. Briefly, given that rates are known, we can infer the distribution of the waiting time to the next event, as well as the probability that the next event is of any particular type. By iterating, we can simulate a stochastic realization.

3.2. Simulating the ensemble distribution. We model the ensemble distribution by creating one conceptual ‘box’ for each possible state of the system, and asking what is the probability that the system is in each box. Conceptually, this approach is surprisingly simple; the catch is that there may be a lot of boxes – in fact, one for each possible state of the system. In an SIS model with fixed

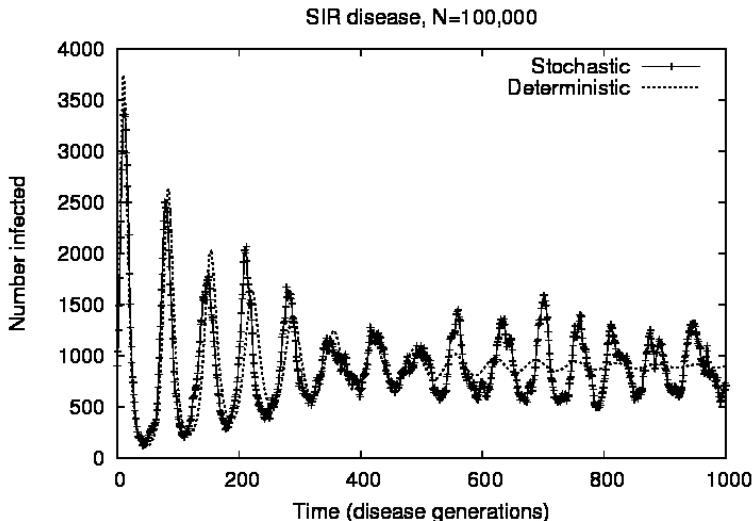


FIGURE 3. The deterministic trajectory from a simple SIR model (2.1) is compared with a typical stochastic trajectory from an analogous stochastic model (Table 1). Parameters are $R_0 = 10$, $\rho = 0.001$, $N = 1,000,000$. The system was started relatively close to the endemic equilibrium (with 120,000 susceptibles) to avoid too violent a first epidemic.

population size, we might have effectively only one state parameter in the deterministic model. In this case, the number probabilities to model in the ensemble distribution is of order N , the population size. In an SIR model, we need at least two deterministic state variables, so the number of discrete state probabilities will be of order N^2 – or N^3 if we allow population size to vary, or take disease latency into account

Since the human global population is on the order of 10^{10} individuals, with major cities having $> 10^5$ individuals, this approach quickly becomes prohibitive. Nonetheless, we may gain valuable information from our investigations of simpler cases.

Although often difficult in practice, simulating an ensemble distribution is conceptually straightforward. Given our event rates, we can determine the rate of *direct* transit between any two states S and S' . This rate will of course be zero in many cases. It is then straightforward to write equations for how the probabilities of being in a given state $p(S)$ change through time:

$$(3.1) \quad \dot{p}_S = \sum_{S'} p_{S'} r_{S' \rightarrow S} - p_S \sum_{S'} r_{S \rightarrow S'}$$

4. Analytic approaches

In addition to simulating equations of the form (3.1), we may also seek analytic understanding. Rather remarkably, in some cases equations of this form can be solved exactly for interesting biological systems. Although these examples are necessarily very limited, such exact solutions can be very illuminating.

We might also seek approximate analytic solutions to increase our understanding. These two approaches are related, since an exact solution to a simplified problem may in some cases serve as an approximate solution to the original problem. In addition to such simplifications, analytic approximation techniques that can be applied include moment approximations and diffusion approximations.

Two of the most useful tools for understanding deterministic disease models are based on linearizations. We linearize around the disease-free equilibrium to investigate what factors control whether the disease can invade and persist. We linearize around the endemic equilibrium to investigate the system's tendency to cycle, and the characteristic period and damping time of cycles. Both of these methods have analogues in demographic models (see below): the linear birth-death model is based on linearizing about the disease-free equilibrium, and diffusion approximations are based on linearizing about the endemic equilibrium.

4.1. The fate of an introduced disease. Consider for a moment the long-term fate of a realization of our demographic system (Table 1). The total population always remains fixed, but on reflection we will see that *eventually* the disease will always go extinct: however large N , and however low the (average) probability of recovery before the next infection, in infinite time, we will always eventually throw H tails in a row, so to speak.

From a biological point of view, this fact may not seem very interesting, since the expected time to extinction of disease in our model will often dwarf the expected lifetime of the universe. Mathematically speaking, however, it poses serious obstacles. Investigating equilibria is a vitally useful mathematical technique, but the only real equilibrium of this system is the disease-free equilibrium. Below, we will see how this concept is addressed with the concept of the quasi-equilibrium.

Beyond the observation that the disease must go extinct eventually, it is useful to categorize qualitatively different *ways* in which it might go extinct. Based on the behavior of our system, we can identify either three (if we are mathematicians) or four (if biologists) such routes to extinction:

- Fizzle::** the disease fails to cause an epidemic. We will make this concept precise below
- Burnout::** the disease causes an epidemic, which reduces the number of susceptibles below the threshold required, and goes extinct before the level of susceptibles recovers (after the first epidemic)
- Fadeout::** the disease causes several epidemics, and approaches “quasi-equilibrium” (see below) before going extinct.
- Persistence::** the disease fades out, but after a biologically unreasonable period of time. There is no convenient *mathematical* distinction between fadeout and persistence, but biologically it may make sense to make an arbitrary distinction.

4.2. Equilibrium and quasi-equilibrium. We define equilibrium in our demographic system as an ensemble distribution that does not change with time. As noted above, the only equilibrium of the system is then the disease-free equilibrium. As long as any populations not extinct, proportion extinct will increase. In particular, there is no equilibrium corresponding to the endemic equilibrium of the deterministic system. This is deeply inconvenient, since we would like to study such an equilibrium, to characterize how the system behaves in the long term.

Event	transition	rate	Effect on I
Infection	$S \rightarrow I$	$R_0 I$	1
Recovery	$I \rightarrow R$	I	-1

TABLE 3. We specify a demographic model by describing the events that can occur; the way that the rates of these events depend on the state of the system; and the effect of these events on the variables that describe the state of our system.

We address the problem by constructing a “quasi-equilibrium”. Consider the ensemble distribution confined to the subset of states where nothing is extinct. This system can be described as a *fully connected* (you can get anywhere from anywhere else, *open* (you can leave the set) flow. Intuition tells us that such a system has a unique, real, negative dominant eigenvalue, associated with a positive eigenvector. Linear algebra tells us this as well [7]. Thus, the system will converge to a stable *relative* distribution of probabilities of being in each non-extinct box – this relative distribution is called the quasi-equilibrium.

The quasi-equilibrium is the asymptotic probability of being in each system state with disease *given that the disease has not gone extinct*. The eigenvalue λ_q associated with the quasi-equilibrium distribution is the rate at which the probability of non-extinction decays (exponentially). The distribution of persistence times must be asymptotically exponential. The expected persistence time (looking forward) approaches $-1/\lambda_q$ as the system continues to persist.

The quasi-equilibrium perspective also gives us a useful tool for numerical investigations. Just as we modeled the ensemble distribution above (3.1), we can also model the *conditional* probability of being in a particular state given that extinction has not occurred. Since these probabilities converge in general to positive values (the quasi-equilibrium distribution), this approach tends to be computationally more stable than directly simulating the ensemble distribution. Since we can also keep track of the cumulative extinction probability, the information we obtain is equivalent. Define the conditional probabilities $q_S = p_S/(1 - p_N)$, where N is a ‘null’ state that we cannot escape from. Then we use the quotient rule to write equations for the conditional probabilities:

$$(4.1) \quad \dot{q}_S = \sum_{S'} q_{S'} r_{S' \rightarrow S} - q_S \sum_{S'} r_{S \rightarrow S'} - q_S r_{S \rightarrow N} + q_S \sum_{S'} q_{S'} r_{S' \rightarrow N},$$

where the sums now go only over the states where the disease is not extinct.

4.3. Linear birth-death process. In deterministic systems, we study disease invasion by asking how the number of infectives behaves in the limit where we assume that virtually the whole population is susceptible. This is (typically) equivalent to linearizing about the disease-free equilibrium. We take an analogous approach with stochastic disease models. This corresponds to holding the number of susceptibles fixed at N ; we are then left with a demographic model whose state is determined by the number of infectious individuals I . This system has only two events:

Surprisingly, the probability of eventual extinction in this system is not one. Our argument that the system must go extinct eventually breaks down, because

the expected number of infectious individuals goes to ∞ in time. Thus, we can ask what is the extinction probability in this model.

This can be solved in a straightforward, and rather illuminating, fashion. Chains of infection are independent in this model: since the rate at which an infection spawns new infections never changes, infectious individuals have no effect on the number of cases that will be caused by another infectious individual. We can use this fact to solve directly for the probability of extinction when starting from I infections, E_I .

If we start with one infectious individual, the probability that the next event will be a recovery (and thus that the disease will go extinct immediately) is $1/(R_0 + 1)$. Otherwise, the next event will be an infection, and we will have two infectious individuals. Thus:

$$(4.2) \quad E_1 = \frac{1 + R_0 E_2}{1 + R_0}.$$

But, if our two new infectious individuals have independent chains of infection, the probability that they will both eventually go extinct is the product of the independent probabilities that each will go extinct: $E_2 = E_1^2$. Thus, we have a quadratic equation for E_1 :

$$(4.3) \quad E_1 = \frac{1 + R_0 E_1^2}{1 + R_0}.$$

This has roots 1 and $1/R_0$, and it can be shown that the eventual extinction probability is $1/R_0$ when this quantity is < 1 , and 1 otherwise. Making use of our independence observation, we conclude that the probability of eventual extinction starting from I infectious individuals is

$$(4.4) \quad E_I = R_0^{-I}, \text{ when } R_0 > 1; 1, \text{ otherwise.}$$

Going back to our anatomy of routes to extinction, we can use E_I as a working definition of fizzle probability: the probability that the disease would have gone extinct even without depleting any susceptibles is a good proxy for the probability of going extinct without causing an epidemic. So this calculation allows us to investigate quite broadly under what circumstances we might expect to see epidemics fizzle. Clearly, if neither R_0 nor I is near 1, we expect E_I to be quite small. Thus, fizzes should be the result either of small introductions, or of very small R_0 ; and when R_0 is not particularly small, we expect epidemics that fizzle to never grow very large.

We derived (4.4) by making use of the fact that (in our model) a single infectious individual will either cause a new, identical infection (increasing I to 2), or recover (decreasing I to 0). It is straightforward to generalize (4.4) to account for the possibility that a single infection will generate any given number of secondary infections, by writing:

$$(4.5) \quad E_1 = \sum_s P_s E_s = \sum_s P_s E_1^s,$$

where P_s is the probability of s secondary cases being generated by a given case, and solving the polynomial for E_1 . We can interpret P_s as reflecting a discrete-time step, for models in discrete time [3], or as reflecting disease generations in either continuous or discrete time, allowing a broad framework that can include parametric heterogeneity [13].

Although this simple calculation provides a great deal of insight, we can in fact investigate our linearized invasion system Table 3 in even more detail, and in fact solve exactly for the ensemble distribution. One way to approach this is through moment methods.

4.4. Moment calculations. Moment methods seek to investigate systems by studying summary statistics called moments: e.g., means, variances, covariances, skewnesses and so on. In the case of our linearized system Table 3, we can simply ask: what is the behavior of the mean, variance, skewness, ..., of I over the ensemble? More complicated systems (such as Table 1) will also require ‘cross’-moments (e.g., covariances).

We start by defining the ensemble mean $\mu = \sum_I I p_I$, and asking how this quantity changes through time. Since the set of possible values of I is fixed (the natural numbers), we can find the change in μ if we know how the p_I change:

$$(4.6) \quad \dot{\mu} = \sum_I I \dot{p}_I.$$

Plugging our event rates (Table 3) into (3.1), we see that we can enter each state two ways (by birth or death), and leave it by the same two ways, so:

$$(4.7) \quad \dot{p}_I = b_{I-1} p_{I-1} + m_{I+1} p_{I+1} - (b_I + m_I) p_I,$$

where $b_I = R_0 I$ is the ‘birth’ rate, and $m_I = I$ is the ‘death’ rate of infectious individuals.

Next, we plug (4.7) into (4.6)

$$(4.8) \quad \dot{\mu} = \sum_I I b_{I-1} p_{I-1} + \sum_I I m_{I+1} p_{I+1} - \sum_I I (b_I + m_I) p_I.$$

Since m_0 , the total death rate when there are no individuals in the population, and p_{-1} , the probability of having -1 individuals in the population, must both be zero, we can re-index our sums to run over p_I (as I runs from 0 to ∞):

$$(4.9) \quad \dot{\mu} = \sum_I (I+1) b_I p_I + \sum_I (I-1) m_I p_I - \sum_I I (b_I + m_I) p_I,$$

which simplifies to

$$(4.10) \quad \dot{\mu} = \sum_I (b_I - m_I) p_I.$$

The result (4.10) is suspiciously simple, and in fact, we could have written it down from first principles: the rate of increase of the mean is equal to the expectation of the birth rate minus the death rate, with the expectation taken over the ensemble distribution. The exercise of formally deriving (4.10) from (3.1) is (I hope) illuminating in several ways, however. It helps shed insight into the ensemble-distribution equations, and gives some guidance for how we might develop moment equations for more complicated systems (or for higher moments).

Note that (4.10), while limited to simple birth-death systems, has not yet made use of our assumption of linearity. Plugging in the rates for our specific linearized birth-death system (Table 3), we find that the ensemble mean obeys:

$$(4.11) \quad \dot{\mu} = \sum_I (R_0 I - I) p_I = (R_0 - 1) \mu$$

We can solve this equation exactly: the expected value of μ (averaged over realizations) grows exponentially through time!

Similar expansions can be made for non-linear systems. In this case, however, we typically see a ‘chain’ of moment equations, where the equation for each moment depends on one or more higher moments. For example, in an SIS model, the per capita “birth” rate will decrease as the number of infectious individuals increases: $b_I = R_0 I(N - I)/N$. Since this model has linear *per capita* rates, we could write an equation for $\dot{\mu}$ that depends on the mean and the variance (i.e., the second moment). The dynamics of the second moment would depend on the third moment, and so on. There are a variety of techniques, broadly called “moment-closure” approximations for ending these chains and approximating the dynamics of such systems.

One approach that is worth mentioning briefly is that of moment-generating functions. This is a nice mathematical trick that involves describing the ensemble distribution as a function of a dummy variable x . In particular, we define $\phi = \sum_I p_I x^I$. If we know the function $\phi(x)$ at a given time, we know all of the probabilities p_I . We can then use our event-rate equations to write an expression for $\dot{\phi}$, which can be converted into a partial differential equation and analyzed, approximated, and (rarely) solved. For example, we can calculate the exact ensemble distribution through time of the linear birth-death process (Table 3). Starting from a fixed number of infectious individuals I_0 , the ensemble distribution is given by a negative binomial distribution with mean $I_0 \exp((R_0 - 1)t)$ and parameter I_0 .

To summarize: we understand the exact behavior of linearized invasion model (Table 3). This allows us to understand the properties of the SIR model, in the absence of effects due to depletion of susceptibles (i.e., its behavior near the disease-free equilibrium). In particular, we can analytically calculate the relationship between model parameters and: the probability that an epidemic will fizzle; the size distribution of fizzled epidemics; and the mean and variance in the expected epidemic growth rate.

4.5. Diffusion approximations. We can approximate the discrete-valued demographic system with a real-valued system that reflects the mean *and variance* of the demographic system [12]. Thus we can incorporate demographic stochasticity in a continuous system. This gives an excellent approximation except when some values are very small. Such systems can be simulated as stochastic differential equations, but if we are going to simulate, we are usually better off simulating the full stochastic system; hybrid approaches also exist, which take advantage of the efficiency of approximating quantities which are not small.

The biggest strength of diffusion approximations is that they allow analytic understanding of demographic stochasticity. In a linear system, we can solve the equilibrium distribution of the continuous equations. We can therefore linearize around the disease-free equilibrium, and use a diffusion approximation to approximate the quasi-equilibrium distribution, and thus ask questions about whether the disease will persist, and how large we expect characteristic demographic fluctuations to be.

We linearize about the endemic equilibrium, in exact analogy to Jacobian methods for stability in deterministic models. A good explanation of the details can be found in [14]. Heuristically speaking, diffusion (and thus demographic stochasticity) is relatively unimportant when the square of the number infected is large compared to the demographic variance. We can ask roughly how we expect these

numbers to scale.

$$(4.12) \quad \begin{aligned} \text{Number infected at equilibrium: } \mu &= \frac{(R_0 - 1)\rho N}{R} \approx \rho N. \\ \text{Demographic variance: } \sigma^2 &\approx N. \end{aligned}$$

We can define a ‘diffusion index’ $U = \mu^2/\sigma^2 \approx \rho^2 N$ [9]. This gives a measure of the size of the infectious population relative to the amount of demographic variability expected. In the linear system, we expect the amount of time required to move a distance of μ from the equilibrium (and thus approach disease extinction), to be very roughly on the order of $\exp(U)$. Conversely, we expect U to give some insight into how large a human population must be to allow the disease to persist for many epidemic cycles (i.e., the critical community size [5])

This very crude calculation gives a major insight into the importance of demographic stochasticity in models of disease with acquired immunity. Recall that the quantity ρ is the ratio between the two key time scales in our system: the duration of infectiousness, and the duration of recalcitrance to infection after recovery. In acute human pathogens, the former is typically days to weeks, and the latter typically years to decades. Thus ρ can easily be as small as 10^{-3} , and demographic stochasticity may have important effects in populations of order $N = 10^6$ or larger. This is in sharp contrast to ecological (or SIS) models, where pure demographic stochasticity is not expected to have large effects in populations larger than a few dozen individual.

It is important to recall the limitations of the linear diffusion approximation in general, and our crude scaling quantity in particular. These methods depend on linearizing about the disease-free equilibrium. Linearization tells us the *local* behavior of system. We expect our diffusion approximation to give a good estimate of the size of small oscillations, but not necessarily of large ones. At parameter ranges of interest, disease extinction typically represents a rather large oscillation: thus, the diffusion approximation will not necessarily give a good estimate of the time to extinction (or of the critical community size).

Put another way, if our diffusion index is small, we expect local behavior, and persistence of the disease (in the sense that expected time to extinction from quasi-equilibrium is ridiculously long). However, if the diffusion index is large, we expect the behavior to become non-local, and thus the index itself is no longer expected to give a reliable guide of whether the disease is likely to go extinct. In particular, and unfortunately, numerical investigations show that the critical community size does not increase as quickly as $1/\rho^2$.

4.6. Conclusions. Treating individuals as individuals can have dramatic effects on models of disease transmission. Although this issue is fundamentally difficult to grapple with, we should give it at least some thought when analyzing, and fitting, models. It is important to recall that populations really are made of individuals, and that analytic results that are not robust to incorporating demographic stochasticity should be regarded with grave suspicion. Stochastic models are fundamentally hard to approach, and we usually combine techniques to understand them; available techniques include analytic approximation, simulating ensemble distributions, and simulating realizations.

Analytic consideration of simple models of disease transmission with demographic stochasticity has led to the development of several useful concepts [2]. In

particular, we can distinguish the phenomena of fizzle, burnout and fadeout, and ask under what circumstances the first two are likely, and are what circumstances the third one is likely to happen within a non-ridiculous time frame. Simple calculations also lead to the conclusion that the relative duration of acquired immunity has extremely strong effects on the likelihood of persistence.

References

- [1] L. J. S. Allen. *An Introduction to Stochastic Processes with Applications to Biology*. Prentice Hall, New Jersey, 2003.
- [2] Roy M. Anderson and Robert M. May. *Infectious Diseases of Humans: Dynamics and Control*. Oxford Science Publications, Oxford, 1991.
- [3] N. J. T. Bailey. *The Mathematical Theory of Infectious Diseases and its Application*. Griffin, London, 1975.
- [4] S. Bansal, B. T. Grenfell, and L. A. Meyers. When individual behaviour matters: homogeneous and network models in epidemiology. *J. R. Soc. Interface*, 4:879–91, 2007.
- [5] M. S. Bartlett. Measles periodicity and community size. *J. Roy. Stat. Soc.*, A120:48–70, 1957.
- [6] M. S. Bartlett. Monte carlo studies in ecology and epidemiology. *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, 4:39–55, 1961.
- [7] A. Berman and R. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, New York, 1979.
- [8] J. Dushoff. Host heterogeneity and disease endemicity: A moment-based approach. *Theor. Popul. Biol.*, 56:325–335, 1999.
- [9] J. Dushoff. Carrying capacity and demographic stochasticity: Scaling behavior of the stochastic logistic model. *Theor. Popul. Biol.*, 57:59–65, 2000.
- [10] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.
- [11] H. W. Hethcote and J. A. Yorke. *Gonorrhea Transmission Dynamics and Control*. Springer-Verlag, Berlin, 1984.
- [12] Egbert G. Leigh, Jr. The average lifetime of a population in a varying environment. *J. Theor. Biol.*, 90:213–239, 1981.
- [13] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz. Superspreading and the effect of individual variation on disease emergence. *Nature*, 438:355–359, 2005.
- [14] I. Nasell. On the time to extinction in recurrent epidemics. *J. R. Stat. Soc. Ser. B-Stat. Methodol.*, 61:309–330, 1999.
- [15] S. Riley. Large-scale spatial-transmission models of infectious disease. *Science*, 316:1298–301, 2007.

DEPARTMENT OF BIOLOGY, MCMASTER UNIVERSITY, HAMILTON, ONTARIO L8S-4K1
 Email address: dushoff@mcmaster.ca