Mathematical Biosciences 261 (2015) 74–82

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Discrete stochastic metapopulation model with arbitrarily distributed infectious period

Nancy Hernandez-Ceron*, Jonathan A. Chavez-Casillas, Zhilan Feng

Purdue University, West Lafayette, IN 47907, United States

ARTICLE INFO

Article history:
Received 11 September 2014
Revised 3 December 2014
Accepted 17 December 2014
Available online 27 December 2014

Keywords:
SIR epidemic model
Probability of minor and major epidemic
Basic reproduction number
Arbitrarily distributed infectious period

ABSTRACT

In this study, a stochastic discrete-time model is developed to study the spread of an infectious disease in an n-patch environment. The model includes an arbitrary distribution of the (random) infectious period T, and the results are used to investigate how the distribution of T may influence the model outcomes. General results are applied to specific distributions including Geometric, Negative Binomial, Poisson and Uniform. The model outcomes are contrasted both numerically and analytically by comparing the corresponding basic reproduction numbers $\mathcal{R}_0$ and probability of a minor epidemic (or probability of disease extinction) $\mathbb{P}_0$. It is shown analytically that for $n = 2$ the reproduction numbers corresponding to different distributions of T can be ordered based on the probability generating function ϕ_T of T. In addition, numerical simulations are carried out to examine the final epidemic size $\mathcal{F}$ and duration of the epidemic $\mathcal{D}$ of a two-patch model.

1. Introduction

Deterministic and stochastic epidemic models have commonly assumed that the disease stages, particularly the infectious period (IP), follow an exponential distribution (continuous-time) or a Geometric distribution (discrete-time). The very property of these distributions that makes these models tractable, the memoryless property, is biologically unrealistic for most infectious diseases. It has been shown that models with these simplifying assumptions may generate misleading assessments on disease control strategies [1,2].

One of the more realistic alternatives to the exponential (Geometric) distribution for the IP that has been considered is the Gamma (Negative Binomial) distribution, which is a natural generalization due to its relationship with the exponential (Geometric) distribution. When a Gamma distribution is considered, the so called "linear chain trick" can be used to reduce the system of integro-differential equations to a system of ordinary differential equations (see, for example, [1,3–5]). The key idea in this approach is to introduce multiple substages for the IP, each of which follows an exponential distribution. A similar idea is applied in stochastic models to allow the use of Gamma distribution for the IP, while still preserving the Markov property of the process. Such models were first developed and studied in [6,7] and more recently in [8,9].

Stochastic models with an arbitrary distribution for the IP were first considered in [10–12], but Sellke's construction [13] helped derive stronger results such as those in [14,15]. Some recent studies have focused on understanding the effect of disease stage distributions on the model outcomes (see, for instance, [16–19]).

In [20], a patch model is used to study the spread of an epidemic through a population divided into n sub-populations (patches), in which individuals move between the patches according to the law of a continuous Markov chain (dynamic population epidemic model). In this framework, infected individuals make contacts with members currently in the same patch. In a more recent study on a continuous-time patch model [21], an expression for the basic reproduction number $\mathcal{R}_0$ and the extinction probability of the epidemic are derived in terms of the IP distribution. It was shown that for a two patch model $\mathcal{R}_0$ is maximized by an IP with constant length. For three or more patches, however, it is very difficult to draw general conclusions about the effects of IP distribution on $\mathcal{R}_0$ or the extinction probability. In the current study, we extend some of the results in [21] to an analogous discrete-time model.

Most epidemic models are in the continuous-time setting, studies on discrete models have been very limited. Mathematical formulations of continuous-time models are in general complicated when an arbitrarily distributed IP is included, particularly when the models also include control measures such as quarantine and isolation (e.g., [1]). This may make it challenging for modelers to communicate with biologists and public health policymakers. Analogous discrete-time models can be formulated in a way that is much easier to understand for non-mathematicians (see, for example, [2,22,23]). Another major

* Corresponding author. Tel.: +1 7654302763.
E-mail addresses: hernann@math.purdue.edu (N. Hernandez-Ceron), jchavezc@math.purdue.edu (J.A. Chavez-Casillas), zfeng@math.purdue.edu (Z. Feng).

http://dx.doi.org/10.1016/j.mbs.2014.12.003
0025-5564/

advantage of discrete-time models is their capability of incorporating distributions directly from empirical data, whereas for continuous-time models one usually needs to estimate the parameters for a standard distribution via data fitting.

In Section 2, the general model with n patches and Markov displacement (with transition matrix D) is described. For an infected individual, the infectious period (T) is assumed to be a discrete random variable with an arbitrary distribution. We derive a formula for the basic reproduction number $\mathcal{R}_0$, which is given by the spectral radius of the mean offspring matrix, a matrix that depends on D and the probability generating function (pgf) of T. An equation for the probability of minor epidemic (extinction probability) $\mathbb{P}_0$ is also derived for this n-patch model.

In Section 3, these general results are then applied to the case $n = 2$ patches. For the two-patch model, in addition to an exact formula, lower and upper bounds for $\mathcal{R}_0$ are also identified. To examine the effect that the distribution of T has on $\mathcal{R}_0$, we consider three specific distributions: shifted Geometric, shifted Negative Binomial, and shifted Poisson. The reproduction numbers corresponding to these distributions have a specific order relation. Numerical simulations for the two-patch model are carried out to explore the influence of the T distribution on the final epidemic size ($\mathcal{F}$), duration of epidemic ($\mathcal{D}$), as well as the probability of disease extinction ($\mathbb{P}_0$).

2. General model

We adopt the approaches used in [20,21] for continuous models to develop a *discrete* stochastic SIR metapopulation model, in a closed population, for an epidemic outbreak with an arbitrarily distribution for the infectious period (IP). The main objective of this study is to investigate how the distribution of IP may affect the model outcomes, particularly the basic reproduction number $\mathcal{R}_0$ and the probability of major epidemic ($1 - \mathbb{P}_0$).

Consider a metapopulation with n sub-populations (patches). Let $N_i(t)$ denote the size of population i at time t for $i = 1, 2, \ldots, n$. Assume that the total population size $N = \sum_{i=1}^{n} N_i(t)$ remains constant for all time. Individuals can move between any two patches, this movement is determined by a **discrete** time Markov chain U, which is described by the transition matrix $D = (\sigma_{ij})$. The entry σ_{ij} represents the probability of moving from population i to population j at each time step.

Effective contacts by individual in population i, per unit of time, is modeled by a Poisson random variable with parameter β_i. In the early stages of the epidemic most effective contacts will produce an infection because most individuals are susceptible. The disease transmission dynamics within each sub-population is governed by an SIR model. It is assumed that individuals become immune after recovery. Let T denote the random variable for the IP (the time until recovery), which is assumed to be the same for all sub-populations. Here, we place no restriction on the T distribution, other than T is discrete, non negative and has a finite mean. All variables and parameters are listed in Table 2. Fig. 1 provides a graphical representation of the model described above.

New infections are produced between time steps in the interval $(t, t + 1)$, while recovery and geographical displacement (governed by the discrete random variables T and U) occur at integer time points. This simplification assumption accompanies discrete models and not their continuous counterpart. However, the assumption is biologically reasonable for different situations, including (i) commuters traveling at peak hours from city to city or (ii) domestic animals who are transported from farm to farm at night.

Assume that, at time $t = 0$, $N_i(0) \approx N\pi_i$ $(i = 1, 2, \ldots, n)$, where $\pi = (\pi_i)_{i=1}^{n}$ is the stationary probability (i.e. $\pi D = \pi$). Thus, although random, the subpopulation $N_i(t)$ will remain close to its initial value throughout time. Some of the properties of the model are described in the following sections.

2.1. Computation of $\mathcal{R}_0$

In this section, we follow the approach of [21]. The early stages of an epidemic is approximated by a properly defined multi-type branching process. The "convergence" of the epidemic model to its associated branching process has been established previously (see [12,24,25]). A less formal but more practical exposition can be found in [9,23,26,27]. To compute the basic reproduction number $\mathcal{R}_0$, we introduce the notation:

$\zeta_{ij} =$ random time spent in patch j (before recovery) by an infectious individual from patch i;
$m_{ij} =$ average number of "offspring" (i.e., secondary infections) that an individual, from patch i, can produce in patch j during the entire "life span" (i.e. the random infectious period modeled by T);
$M = (m_{ij})$, the mean offspring matrix.

$\mathcal{R}_0$ is given by the spectral radius of the matrix M, which entries m_{ij} can be written as

$$m_{ij} = \beta_j \mathbb{E}(\zeta_{ij}). \tag{1}$$

By conditional expectation $\mathbb{E}(\zeta_{ij}) = \sum_{t=0}^{\infty} \mathbb{E}(\zeta_{ij}|T = t)\mathbb{P}(T = t)$ and

$$\mathbb{E}(\zeta_{ij}|T = t) = \mathbb{E}\left(\sum_{k=0}^{t-1} \mathbb{I}_{U_i(k)=j}\right) = \sum_{k=0}^{t-1} \mathbb{P}(U_i(k) = j) = \sum_{k=0}^{t-1} \sigma_{ij}^{(k)},$$

where $\sigma_{ij}^{(k)}$ denotes the ijth entry of the matrix D^k, $U_i(k)$ the state of the Markov chain at time k given that $U_i(0) = i$, and $\mathbb{I}_{U_i(k)=j}$ the indicator function of the event $U_i(k) = j$. Notice that new infections at time t are generated by infective individuals at time $t - 1$, which is why the sum above has been taken from 0 to $t - 1$. Combining the last two equations we obtain the matrix of expectations of ζ_{ij}

$$\begin{bmatrix} \mathbb{E}(\zeta_{11}) & \cdots & \mathbb{E}(\zeta_{1n}) \\ \vdots & \ddots & \vdots \\ \mathbb{E}(\zeta_{n1}) & \cdots & \mathbb{E}(\zeta_{nn}) \end{bmatrix} = \sum_{t=1}^{\infty} \mathbb{P}(T = t) \sum_{k=0}^{t-1} D^k = \mathbb{E}\left(\sum_{k=0}^{T-1} D^k\right). \tag{2}$$

Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of the stochastic matrix $D = (\sigma_{ij})$. Since D is a Markov matrix, $\lambda_i = 1$ for some i and $|\lambda_i| \leq 1$ $\forall i$. If D is

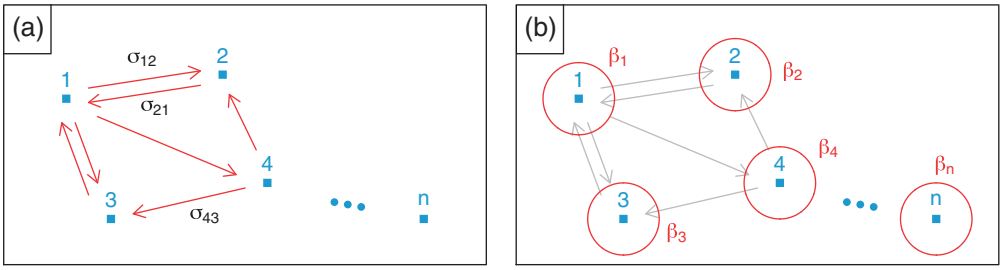

Fig. 1. (a) Individuals move from patch to patch at time $t \in \mathbb{N}$ according to the Markov chain U. (b) Once the infection process has started in one patch, the disease can spread to other patches. Contacts by an infected individual, per unit of time in patch i, is described by Poisson(β_i).