Mathematical Biosciences 228 (2010) 45–55

Contents lists available at ScienceDirect

Mathematical Biosciences

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

On the number of recovered individuals in the *SIS* and *SIR* stochastic epidemic models

J.R. Artalejo [a,*], A. Economou [b], M.J. Lopez-Herrero [c]

[a] Department of Statistics and Operations Research, Faculty of Mathematics, Complutense University of Madrid, 28040 Madrid, Spain
[b] Department of Mathematics, University of Athens, Panepistemiopolis, 15784 Athens, Greece
[c] School of Statistics, Complutense University of Madrid, 28040 Madrid, Spain

ARTICLE INFO

Article history:
Received 21 June 2009
Received in revised form 6 August 2010
Accepted 21 August 2010
Available online 27 August 2010

Keywords:
Recovered individuals
Stochastic *SIS* model
Stochastic *SIR* model
Extinction time
Transient behavior
Outbreak of *ESBL*

ABSTRACT

The basic models of infectious disease dynamics (the *SIS* and *SIR* models) are considered. Particular attention is paid to the number of infected individuals that recovered and its relationship with the final epidemic size. We investigate this descriptor both until the extinction of the epidemic and in transient regime. Simple and efficient methods to obtain the distribution of the number of recovered individuals and its moments are proposed and discussed with respect to the previous work. The methodology could also be extended to other stochastic epidemic models. The theory is illustrated by numerical experiments, which demonstrate that the proposed computational methods can be applied efficiently. In particular, we use the distribution of the number of individuals removed in the *SIR* model in conjunction with data of outbreaks of *ESBL* observed in the intensive care unit of a Spanish hospital.

1. Introduction

There exists a vast literature that employs Markov chains for modelling the stochastic dynamics of an epidemic, see, for example, [2,3,8,9,13,15,17,18,21,22,29,34]. In this context, the interest is mainly focused on the distribution of the epidemic size (i.e., quasi-stationary distribution, final size distribution) and the extinction time.

In many models, there is no external source of infection so the spread of the epidemic ends at the first time at which all the infected individuals are recovered. In this case, the Markov chain describing the epidemic size has a set of absorbing states. As a result, the stationary distribution is degenerate. However, when the time to absorption is sufficiently long, one may deal with the quasi-stationary distribution which provides a measure of the system dynamics conditioning on the event that the absorption has not occurred yet. On the other hand, the risk of extinction is typically measured in terms of the extinction probability (i.e., the probability of reaching the absorption in a finite time) and the expected time to extinction. Since the extinction time is the time to hit the set of absorbing states, it gives a measure of a continuous nature.

Some models that assume an external source of infection (e.g., immigration, outside force, environmental indirect transmission, reintroduction parameter, etc.) [15,28,29,33] provide motivation to eliminate the absorbing states. Then, the stationary distribution can be investigated. Moreover, the length of an individual outbreak amounts to the study of the first absence of infection.

Our aim in this paper is to study a natural discrete counterpart of the extinction time namely the number of recovered individuals. For the susceptible-infected-removed (*SIR*) stochastic model (see also literature regarding the so-called general stochastic epidemic), the number of recovered individuals amounts to the final epidemic size, provided that the study is performed over the entire duration of the epidemic. In this context, the early literature includes at least two approaches due to Bailey [7] and Whittle [32] for the recursive solution of the triangular systems of equations governing the final size distribution of the *SIR* model. Both approaches are also summarized in the book [8]. The reader is also referred to [18,20,21,26] where other alternative methods are presented. These methods are based on a variety of techniques including combination of Laplace transforms and a matrix approach [18], spectral approach based on the eigenvectors of the infinitesimal generator [21] and a recursive algorithm performed on the embedded epochs when an event occurs [20,26]. Ball [10] generalized the study to cover the case of arbitrary infectious period distribution. Later

* Corresponding author.
E-mail addresses: jesus_artalejo@mat.ucm.es (J.R. Artalejo), aeconom@math.uoa.gr (A. Economou), lherrero@estad.ucm.es (M.J. Lopez-Herrero).

0025-5564/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.mbs.2010.08.006

Demeris and O'Neill [19] used multiple precision arithmetic for overcoming the numerical problems arising in practical computations. There also exists an important number of papers where the study of the final size is extended to more sophisticated but realistic epidemic models including households, multipopulation and rumours; see, for instance, [10,12,20,25] and references therein. As a few selected examples of approximations and asymptotic results, we mention the papers [11,13,25].

Our purpose in this paper is twofold. First, we are interested in the distribution of the number of recovered individuals, not only during the time until the absorption, but also in transient regime. This objective is a distinguished feature of our approach in comparison with the existing literature which is referred to the entire time till absorption. In our opinion, both approaches are worthy of being investigated. First, the analysis of the descriptor till the end of the epidemic provides a clear picture of the global magnitude of the epidemic. However, sometimes the time to extinction may be extremely long and it may be useful to take management actions in the light of the information collected during a time interval, let us say $[0,t]$. Thus, the transient approach is also useful in its own right. As related work, we mention the recent paper [5], where the transient approach was used to study the maximum number of infected individuals in the *SIS* model.

The second purpose of the paper is to study the number of recovered individuals in the susceptible-infected-susceptible (*SIS*) stochastic model. In the *SIS* epidemic model, the set of absorbing states reduces to the sole state 0. Moreover, individuals can recover from the disease but do not develop immunity. Thus, a marked individual can be infected, and consequently recovered, several times. This indicates that the number of recovered individuals (i.e., the subject matter of this paper) is different from the final epidemic size defined for *SIR* models, but it gives a natural analog. Similar descriptors are independently considered in [4], for a population growth model, and in [29], where the size of an individual outbreak is measured in terms of the number of infection cases. Later on, in Section 2, we will show how the number of recovered individuals and the number of infected individuals determine each other reciprocally. Except these two recent papers, as far as the authors know, the number of infected individuals that recovered has not been investigated in the *SIS* setting.

Along the paper, we focus on the basic formulations of the classical *SIS* and *SIR* epidemic models as they can be found in many textbooks [2,9,18]. A replicated version of the *SIR* model provides the appropriate framework for the investigation of the outbreaks of extended-spectrum beta-lactamase (*ESBL*) reported from the Fundación Jiménez Díaz – Capio hospital. Our methodology could be extended to more complicated stochastic epidemic models. In this sense, we mention that the existing literature is rich in variants and generalizations, please see some recent papers [12,15,22,29,34] and references therein.

The rest of the paper is organized as follows. In Section 2, we consider the *SIS* model. We perform an exhaustive analysis of the number of recovered individuals both before the extinction and in transient regime. Our study includes generating functions, moments and a direct computational method for the mass probabilities. A parallel analysis for the *SIR* model is given in Section 3. Our numerical experiments, including the investigation of the outbreaks of *ESBL*, are presented in Section 4. Concluding remarks are presented in Section 5. For the sake of completeness, we also give results for the transient study of the number of infected individuals. Since the arguments are similar to those given for the number of recovered individuals, we avoid unnecessary repetitions and we just summarize the main results in the Appendices.

2. The basic *SIS* epidemic model

The model we consider is the stochastic *SIS* model. However, we first describe a more general birth-and-death process. Consider a closed population of size N. At time t, the population consists of $I(t)$ infected individuals and $S(t) = N - I(t)$ susceptible individuals. We suppose that the process $\{I(t);\ t \geqslant 0\}$ is a birth-and-death process with state space $S = \{0,\ldots,N\}$. Given that there are i infected individuals, the birth rates $\lambda_i > 0$, for $1 \leqslant i \leqslant N-1$, ($\lambda_N = 0$) correspond to transitions occurring when a susceptible individual becomes infected, whereas the death rates $\mu_i > 0$, for $1 \leqslant i \leqslant N$, are associated to the recovery of infected individuals. From this description, it is assumed that individuals do not develop immunity whenever they recover. The epidemic ends as soon as there are no infectives in the population. Thus, we take $\lambda_0 = 0$ in agreement with the fact that the origin is an absorbing state. Since $S\backslash\{0\}$ is finite and irreducible, the absorption occurs in a finite time with probability one.

From the general birth-and-death formulation, we may derive many particular cases including the basic formulation of the stochastic *SIS* model, whose rates are given by $\lambda_i = \frac{\beta}{N} i(N-i)$ and $\mu_i = \gamma i$. The parameters β and γ denote the effective contact and recovery rates, respectively. The transitions among states are represented in Fig. 1.

2.1. The number of individuals recovered before the extinction

We are concerned with the distribution of the recovered individuals prior to disease extinction. Given the initial condition $I(0) = i$, we denote by N_i^R the total number of individuals recovered before the extinction. We further denote the number of individuals infected during this period by N_i^I. Then, it is clear that

$$N_i^R = N_i^I + i. \tag{2.1}$$

Since N_i^R and N_i^I determine each other through formula (2.1), we only need to study one of them, let us say N_i^R. In what follows, we omit the superscript and denote N_i^R simply by N_i.

In the rest of this section, we study the main characteristics of the distribution of N_i. An alternative transient analysis which is referred to any time interval $[0,t]$ rather than to the extinction time is presented in Section 2.2.

2.1.1. The generating functions

Let $\phi_i(z)$ be the generating function of N_i; that is, $\phi_i(z) = E[z^{N_i}] = \sum_{k=i}^{\infty} z^k P\{N_i = k\}$, for $|z| \leqslant 1$. By conditioning on the state visited after the first transition, we get the equations governing the generating functions $\phi_i(z)$:

$$\phi_0(z) = 1, \tag{2.2}$$

$$\phi_i(z) = \frac{z\mu_i}{\lambda_i + \mu_i}\phi_{i-1}(z) + \frac{\lambda_i}{\lambda_i + \mu_i}\phi_{i+1}(z), \quad 1 \leqslant i \leqslant N. \tag{2.3}$$

From Eqs. (2.2) and (2.3), it is possible to develop a stable recursive scheme involving only algebraic operations with positive terms. The trick is to use a forward-elimination-backward-substitution (*FEBS*) argument in combination with an appropriate change of variables. For details, see [5] where similar schemes are also obtained. Then, numerical inversion can be performed with the help of a Fast Fourier Transform (*FFT*) algorithm [30]. In this way, we have a first method for the computation of the

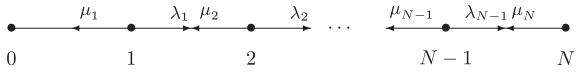

Fig. 1. States and transitions in the *SIS* epidemic model.