A TIME DEPENDENT SIMPLE STOCHASTIC EPIDEMIC

GRACE L. YANG
UNIVERSITY OF MARYLAND
and
CHIN LONG CHIANG
UNIVERSITY OF CALIFORNIA, BERKELEY

1. Introduction

Since the pioneer work of A. M. McKendrick in 1926, many authors have contributed to the advancement of the stochastic theory of epidemics, including Bartlett [4], Bailey [1], D. G. Kendall [12], Neyman and Scott [13], Whittle [16], to name a few. Mathematical complexity involved in some of the epidemic models has aroused the interest of many others. For example, the general stochastic epidemic model where a population consists of susceptibles, infectives, and immunes (see [2], p. 39), has motivated Kendall to suggest an ingenious device. Other authors also have investigated various aspects of the problem. (See, for example, Daniels [8], Downton [9], Gani [11] and Siskind [15].) The model discussed in the present paper deals with a closed population without removal of infectives, a special case of which has been studied very extensively by Bailey [3]. Following Bailey, we label it "a time dependent simple stochastic epidemic."

In a simple stochastic epidemic model, a population consists of two groups of individuals: susceptibles and infectives; there are no removals, no deaths, no immunes, and no recoveries from infection. At the initial time $t = 0$, there are N susceptibles and 1 infective. For each time t, for $t > 0$, there are a number of infectives denoted by $Y(t)$ and a number of uninfected susceptibles $X(t)$, with $Y(t) + X(t) = N + 1$, the total population size remaining unchanged. Our primary purpose is to derive an explicit solution for the probability distribution of the random variable $Y(t)$,

$$P_{1n}(0, t) = Pr\{Y(t) = n | Y(0) = 1\}, \qquad n = 1, \cdots, N + 1. \tag{1}$$

For each interval (τ, t), $0 \leqq \tau \leqq t < \infty$, and for each n, we assume the existence of a nonnegative continuous function $\beta_n(\tau)$ such that

$$\frac{\partial}{\partial t} P_{n,m}(\tau, t)\Big|_{t=\tau} = \begin{cases} -\beta_n(\tau) & \text{for } m = n, \\ \beta_n(\tau) & \text{for } m = n + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Under the assumption of homogeneous mixing of the population, we let

$$\beta_n(\tau) = n(N + 1 - n)\beta(\tau) = a_n\beta(\tau), \tag{3}$$

where

$$a_n = n(N + 1 - n). \tag{4}$$

The quantity $\beta(\tau)$, which is a function of time τ, is known as the infection rate. Thus, in this model, the intensity of spreading of disease may vary with time during an epidemic. It follows from (2) that, for each $t > 0$, the probability function $P_{1n}(0, t)$ satisfies the following system of differential equations

$$\begin{aligned} \frac{d}{dt} P_{11}(0, t) &= -a_1\beta(t)P_{11}(0, t) \\ \frac{d}{dt} P_{1n}(0, t) &= -a_n\beta(t)P_{1n}(0, t) + a_{n-1}\beta(t)P_{1,n-1}(0, t) \end{aligned} \tag{5}$$

for $n = 2, \cdots, N + 1$, with the initial condition $P_{11}(0, 0) = 1$.

Equations (5) are essentially the same as those studied extensively by Bailey [1], [2], [3], except that in those publications the infection rate is assumed to be independent of time (that is, $\beta(t) = \beta$) and the random variable is $X(t)$, the number of susceptibles remaining at time t. Bailey used the Laplace transform, the generating function, and a very skillful mathematical manipulation to provide the solution. However, the computations involved are too complex. Yang has recently established a relationship between the density function of the time of infections and the probability of the number of infections to arrive at a solution [17]. In the present paper, we offer another approach to the problem.

The present solution of system (5) requires the following two lemmas.

LEMMA 1. *Whatever may be distinct real numbers* $a_1, \cdots, a_n$,

$$\sum_{i=1}^{n} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{n} (a_i - a_\alpha)} = 0. \tag{6}$$

Lemma 1 may be found in Pólya and Szegö [14]. Several proofs of the lemma have been given in Chiang [5], [6], pp. 126–127.

LEMMA 2. *Whatever may be* k, *for* $1 \leqq k < n$, *the probabilities in* (1) *satisfy the equality*

$$P_{1n}(0, t) = \int_0^t P_{1k}(0, \tau)a_k\beta(\tau)P_{k+1,n}(\tau, t)\, d\tau. \tag{7}$$

Equation (7) may be easily justified. Let k be an arbitrary but *fixed* integer, $1 \leqq k < n$; the $(k + 1)$th infection must take place somewhere between 0 and t. Let it take place in interval $(\tau, \tau + d\tau)$; then there are k infectives at τ, and $(n - k - 1)$ infectives occurring during (τ, t); the corresponding probability is

$$P_{1k}(0, \tau)a_k\beta(\tau)\, d\tau\, P_{k+1,n}(\tau, t), \tag{8}$$

where $P_{k+1,n}(\tau, t)$ is the conditional probability of n infectives at t given $k + 1$ infectives at τ. Since the events corresponding to the probability (8) for different τ are mutually exclusive, we may integrate (8) from $\tau = 0$ to $\tau = t$ to obtain the required equation (7). Equation (7) holds true whatever may be $1 \leqq k < n$ and regardless of whether the a_i are distinct. For a general discussion on the lemma, the reader is referred to Chiang [7].

2. Solution for the probability $P_{1n}(0, t)$

Solution of the differential equations in (5) depends on whether $n \leqq (N+1)/2$ or $n > (N+1)/2$. The two cases are presented separately below.

Case 1: $1 \leqq n \leqq (N+1)/2$. For these values of n, $a_1, \cdots, a_n$ are all distinct; the differential equations in (5) have the solution

$$(9)\quad P_{1n}(0, t) = (-1)^{n-1} a_1 \cdots a_{n-1} \left[\sum_{i=1}^{n} \frac{\exp\{-a_i \lambda(t)\}}{\prod_{\alpha=1, \alpha \neq i}^{n} (a_i - a_\alpha)} \right], \qquad n = 1, \cdots, \frac{N}{2} \text{ or } \frac{N+1}{2},$$

where

$$(10)\quad \lambda(t) = \int_0^t \beta(\tau)\, d\tau$$

is assumed to be such that $\lim_{t \to \infty} \lambda(t) = \infty$. We assume that $a_0 = 1$ and $\pi(a_i = a_\alpha) = 1$ for $n = 1$.

Solution (9), which can be verified by induction using Lemma 1, is similar to that in the pure birth process (see, for example, Feller [10] and Chiang [6], pp. 51–52), except that, in the present case, $\beta(\tau)$ is a function of time.

When $\beta(\tau) = \beta$, $\lambda(t) = \beta t$, and solution (9) becomes

$$(11)\quad P_{1n}(0, t) = (-1)^{n-1} a_1 \cdots a_{n-1} \left[\sum_{i=1}^{n} \frac{\exp\{-a_i \beta t\}}{\prod_{\alpha=1, \alpha \neq i}^{n} (a_i - a_\alpha)} \right], \qquad n = 1, \cdots, \frac{N}{2} \text{ or } \frac{N+1}{2}.$$

For a_i defined in (4),

$$(12)\quad a_1 \cdots a_{n-1} = (n-1)! \frac{N!}{(N+1-n)!},$$

$$(13)\quad \prod_{\alpha=1, \alpha \neq i}^{n} (a_i - a_\alpha) = (-1)^{n-i} \frac{(i-1)!(n-i)!(N-i)!}{(N-2i+1)(N-i-n)!},$$

and solution (11) may be rewritten

$$(14)\quad P_{1n}(0, t) = \sum_{i=1}^{n} (-1)^{i-1} \frac{(N-2i+1)(n-1)!N!(N-i-n)! \exp\{-a_i \beta t\}}{(i-1)!(N-i)!(N+1-n)!(n-i)!},$$

for $n = 1, 2, \cdots, (N/2$ or $(N+1)/2)$, which is the same as that obtained by Bailey [2].

Case 2: $(N+1)/2 < n \leqq N+1$. Formula (9) no longer holds true when $n > (N+1)/2$ for the reason that in this case the a_i are not all distinct, and in particular,

$$a_i = i(N + 1 - i) = a_{N+1-i}. \tag{15}$$

However, solution of the differential equations in (5) can be obtained by using Lemma 2. In applying equality (7) to the present problem, the integer k must be chosen so that the a_i in the probability $P_{1k}(0, \tau)$ are distinct and the a_i in $P_{k+1,n}(\tau, t)$ also are distinct. When N is even, $k = N/2$; when N is odd, $k = (N + 1)/2$.

With these values of k, we apply formula (9) to the two probabilities in the integrand in equation (7) to obtain

$$P_{1k}(0, \tau) = (-1)^{k-1}a_1 \cdots a_{k-1}\left[\sum_{i=1}^{k} \frac{\exp\{-a_i\lambda(\tau)\}}{\prod_{\alpha=1,\alpha\neq i}^{k}(a_i - a_\alpha)}\right] \tag{16}$$

and

$$P_{k+1,n}(\tau, t) = (-1)^{n-k-1}a_{k+1} \cdots a_{n-1}\left[\sum_{j=k+1}^{n} \frac{\exp\{-a_j[\lambda(t) - \lambda(\tau)]\}}{\prod_{\delta=k+1,\delta\neq j}^{n}(a_j - a_\delta)}\right]. \tag{17}$$

Substituting (16) and (17) in (7) gives the basic formula

$$P_{1n}(0, t) = (-1)^n a_1 \cdots a_{n-1} \sum_{i=1}^{k}\sum_{j=k+1}^{n}\int_0^t \frac{\exp\{-a_i\lambda(\tau)\}\exp\{-a_j[\lambda(t) - \lambda(\tau)]\}}{\prod_{\alpha=1,\alpha\neq i}^{k}(a_i - a_\alpha)\prod_{\delta=k+1,\delta\neq j}^{n}(a_j - a_\delta)}\beta(\tau)\,d\tau. \tag{18}$$

The integral in (18) depends on the values of a_i and a_j. According to the definition of $\lambda(t)$ in (10),

$$\int_0^t \exp\{-a_i\lambda(\tau)\}\exp\{-a_j[\lambda(t) - \lambda(\tau)]\}\beta(\tau)\,d\tau = \frac{-1}{a_i - a_j}[\exp\{-a_i\lambda(t)\} - \exp\{-a_j\lambda(t)\}], \quad a_i \neq a_j, \tag{19}$$

and

$$\int_0^t \exp\{-a_i\lambda(\tau)\}\exp\{-a_j[\lambda(t) - \lambda(\tau)]\}\beta(\tau)\,d\tau = \lambda(t)\exp\{-a_i\lambda(t)\}, \quad a_j = a_i. \tag{20}$$

There are $(n - k)$ terms where $a_i = a_j$ with $i + j = 2k + 1$ when $N = 2k$, and $i + j = 2k$ when $N = 2k - 1$; they are

$$a_{2k+1-n} = a_n, a_{2k+2-n} = a_{n-1}, \cdots, a_k = a_{k+1}, \tag{21}$$

for $N = 2k$, and

$$a_{2k-n} = a_n, a_{2k+2-n} = a_{n-1}, \cdots, a_{k-1} = a_{k+1}, \tag{22}$$

for $N = 2k - 1$. The probabilities $P_{1n}(0, t)$ assume slightly different forms for $N = 2k$ and for $N = 2k - 1$.

(i) *N is even: $N = 2k$.* Substituting (19) and (20) in (18) gives the desired formula for the probability

$$(23)\quad P_{1n}(0,t) = (-1)^{n-1}a_1 \cdots a_{n-1}\left[-\sum_{i=2k+1-n}^{k} \frac{\lambda(t)\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{n}(a_i-a_\alpha)} + \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{n} \frac{\exp\{-a_i\lambda(t)\}-\exp\{-a_j\lambda(t)\}}{(a_i-a_j)\prod_{\alpha=1,\alpha\neq i}^{k}(a_i-a_\alpha)\prod_{\delta=k+1,\delta\neq j}^{n}(a_j-a_\delta)}\right],$$

for $n = k+1, \cdots, N$, where $k = N/2$.

Note that in the product $\prod_{\alpha=1}^{n}(a_i - a_\alpha)$ in formula (23) there are two values of α for which $a_\alpha = a_i$; namely, a_i and a_{N+1-i}; they are both excluded from the product.

The probability $P_{1,N+1}(0,t)$ may be computed from

$$(24)\quad P_{1,N+1}(0,t) = \int_0^t P_{1,N}(0,\tau)a_N\beta(\tau)\,d\tau$$

$$= a_1\cdots a_N\left[\sum_{i=1}^{k}\frac{\int_0^t \lambda(\tau)\exp\{-a_i\lambda(\tau)\}\beta(\tau)\,d\tau}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i-a_\alpha)} - \sum_{i=1}^{k}\sum_{j=k+1}^{N}\frac{\int_0^t[\exp\{-a_i\lambda(\tau)\}-\exp\{-a_j\lambda(\tau)\}]\beta(\tau)\,d\tau}{\prod_{\alpha=1}^{k}(a_i-a_\alpha)\prod_{\delta=k+1}^{N}(a_j-a_\delta)(a_i-a_j)}\right].$$

The first integral of (24) is evaluated to give

$$(25)\quad \int_0^t \lambda(\tau)\exp\{-a_i\lambda(\tau)\}\beta(\tau)\,d\tau = \frac{1}{a_i^2} - \frac{\exp\{-a_i\lambda(t)\}}{a_i^2} - \frac{\lambda(t)\exp\{-a_i\lambda(t)\}}{a_i}.$$

Thus, the first term inside the brackets in (24) becomes

(26)

$$\sum_{i=1}^{k}\frac{1}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i-a_\alpha)a_i^2} - \sum_{i=1}^{k}\frac{\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i-a_\alpha)a_i^2} - \sum_{i=1}^{k}\frac{\lambda(t)\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i-a_\alpha)a_i}.$$

The second integral in (24) is

$$(27)\quad \int_0^t[\exp\{-a_i\lambda(\tau)\}-\exp\{-a_j\lambda(\tau)\}]\beta(\tau)\,d\tau = -\frac{a_i-a_j}{a_ia_j} - \left[\frac{\exp\{-a_i\lambda(t)\}}{a_i} - \frac{\exp\{-a_j\lambda(t)\}}{a_j}\right],$$

and the second term inside the brackets in (24) becomes

$$(28)\quad \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N}\frac{1}{\prod_{\alpha=1}^{k}(a_i-a_\alpha)\prod_{\delta=k+1}^{N}(a_j-a_\delta)a_ia_j} + \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N}\frac{[\exp\{-a_i\lambda(t)\}/a_i]-[\exp\{-a_j\lambda(t)\}/a_j]}{\prod_{\alpha=1,\alpha\neq i}^{k}(a_i-a_\alpha)\prod_{\delta=k+1,\delta\neq j}^{N}(a_j-a_\delta)(a_i-a_j)}$$

Combining the two constant terms in (26) and (28), and using Lemma 1, we have

$$(29)\quad \sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N} (a_i - a_\alpha)a_i^2} + \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N} \frac{1}{\prod_{\alpha=1}^{k}(a_i - a_\alpha)\prod_{\delta=k+1}^{N}(a_j - a_\delta)a_i a_j}$$

$$= \left[\sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1}^{k}(a_i - a_\alpha)a_i}\right]^2 = \left[\frac{1}{\prod_{i=1}^{k}(-a_i)}\right]^2 = \frac{1}{\prod_{i=1}^{N} a_i}.$$

In the second term in (28), the running indices i and j are interchangeable, so that

$$(30)\quad \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N} \frac{[\exp\{-a_i\lambda(t)\}/a_i] - [\exp\{-a_j\lambda(t)\}/a_j]}{\prod_{\alpha=1}^{k}(a_i - a_\alpha)\prod_{\delta=k+1}^{N}(a_j - a_\delta)(a_i - a_j)}$$

$$= 2\sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N} \frac{\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1}^{k}(a_i - a_\alpha)\prod_{\delta=k+1}^{N}(a_j - a_\delta)(a_i - a_j)a_i}.$$

With the simplifications in (29) and (30), we substitute (26) and (28) in (24) to obtain the formula

$$(31)\quad P_{1,N+1}(0, t)$$

$$= 1 - a_1 \cdots a_N\left[\sum_{i=1}^{k} \frac{\lambda(t)\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i - a_\alpha)a_i} + \sum_{i=1}^{k} \frac{\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N}(a_i - a_\alpha)a_i^2}\right.$$

$$\left. - 2\sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{N} \frac{\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,\alpha\neq i}^{k}(a_i - a_\alpha)\prod_{\delta=k+1,\delta\neq j}^{N}(a_j - a_\delta)(a_i - a_j)a_i}\right],$$

where $k = N/2$ and $\lambda(t) = \int_0^t \beta(\tau)\, d\tau$.

(ii) *N is odd:* $N = 2k - 1$. The essential difference between this case and the preceding one is in the limits of the summations and the value of a_k (that is, $a_{(N+1)/2}$) which is now distinct from all other a_i. Keeping these differences in mind, we again substitute (19) and (20) in (18) to obtain the probabilities

$$(32)\quad P_{1n}(0, t) = (-1)^{n-1}a_1 \cdots a_{n-1}\left[-\sum_{i=2k-n}^{k-1} \frac{\lambda(t)\exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{n}(a_i - a_\alpha)}\right.$$

$$\left. + \sum_{i=1}^{k}\sum_{\substack{j=k+1\\ a_i\neq a_j}}^{n} \frac{\exp\{-a_i\lambda(t)\} - \exp\{-a_j\lambda(t)\}}{\prod_{\alpha=1,\alpha\neq i}^{k}(a_i - a_\alpha)\prod_{\delta=k+1,\delta\neq j}^{n}(a_j - a_\delta)(a_i - a_j)}\right],$$

for $n = k + 1, \cdots, N$; with $k = (N + 1)/2$, and

$$(33)\quad P_{1,N+1}(0,t) = 1 - a_1 \cdots a_N \left[\sum_{i=1}^{k-1} \frac{(\lambda(t) + a_i^{-1}) \exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1}^{N} (a_i - a_\alpha) a_i} + \sum_{i=1}^{k} \sum_{\substack{j=k+1 \\ a_i \neq a_j}}^{N} \frac{[\exp\{-a_i\lambda(t)\}/a_i] - [\exp\{-a_j\lambda(t)\}/a_j]}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha) \prod_{\delta=k+1,\delta\neq j}^{N} (a_i - a_\delta)(a_i - a_j)} \right].$$

In formulas (9), (23), and (32) of the probabilities $P_{1n}(0, t)$, every term contains a factor $\exp\{-a_i\lambda(t)\}$ with $a_i > 0$. Therefore, as $t \to \infty$, $P_{1n}(0, t) \to 0$, for $n = 1, \cdots, N$; whereas formulas (31) and (33) show that $P_{1,N+1}(0, t) \to 1$ as $t \to \infty$. This means that in the simple epidemic model considered here, all the N susceptibles will be infected sooner or later; and the epidemic is said to be complete (see Bailey [2]).

3. Infection time and duration of the epidemic

The length of time elapsed till the occurrence of the nth infection is a continuous random variable taking on nonnegative real numbers. Let it be denoted by T_n, for $1 < n \leq N + 1$, with $T_1 = 0$. When $n = N + 1$, T_{N+1} is the duration of the epidemic. The purpose of this section is to derive explicit formulas for the density $f_n(t)$, the distribution function $F_n(t)$, the expectation and variance of T_n.

The density function $f_n(t)$ has a close relationship with the probability $P_{1,n-1}(0, t)$ of $n - 1$ infectives at time t. By definition, $f_n(t)\,dt$ is the probability that the random variable T_n will assume values in the interval $(t, t + dt)$. This means that at time t there are $n - 1$ infectives and the nth infection takes place in interval $(t, t + dt)$; the probability of the occurrence of these events is $P_{1,n-1}(0, t)a_{n-1}\beta(t)\,dt$. Therefore, we have the density function

$$(34)\qquad f_n(t)\,dt = P_{1,n-1}(0, t)a_{n-1}\beta(t)\,dt,$$

and, hence, the distribution function

$$(35)\qquad F_n(t) = \int_0^t P_{1,n-1}(0, \tau)a_{n-1}\beta(\tau)\,d\tau, \qquad n = 2, \cdots, N.$$

Using the formulas of the probabilities $P_{1,n-1}(0, t)$ in the preceding section, we can write down explicit functions for $f_n(t)$ and $F_n(t)$ for each n. We give two examples below.

EXAMPLE 1: $n \leq (N + 1)/2$. We substitute formula (9) in (34) and (35) to obtain the density function

$$(36)\qquad f_n(t)\,dt = (-1)^{n-2}a_1 \cdots a_{n-1} \left[\sum_{i=1}^{n-1} \frac{\beta(t) \exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,\alpha\neq i}^{n-1} (a_i - a_\alpha)} \right] dt,$$

and the distribution function

$$(37)\qquad F_n(t) = (-1)^{n-2}a_1 \cdots a_{n-1} \sum_{i=1}^{n-1} \frac{1 - \exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,\alpha\neq i}^{n-1} (a_i - a_\alpha)a_i}$$

for $n = 2, \cdots, N/2$ or $(N+1)/2$, and $0 < t < \infty$. As $t \to \infty$, $f_n(t) \to 0$ and

$$F_n(\infty) = (-1)^{n-2} a_1 \cdots a_{n-1} \sum_{i=1}^{n-1} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{n-1} (a_i - a_\alpha) a_i} = 1, \tag{38}$$

since Lemma 1 implies that

$$\sum_{i=1}^{n-1} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{n-1} (a_i - a_\alpha) a_i} = -\frac{1}{\prod_{i=1}^{n-1} (-a_i)}. \tag{39}$$

EXAMPLE 2: *the duration of epidemic* T_{N+1}, *when* $N = 2k$. In this case formula (23) for $n = N$ is used in (34) and (35). The density function and the distribution function for T_{N+1} are, respectively,

$$\begin{aligned} f_{N+1}(t)\,dt = (-1)a_1 \cdots a_N \Bigg[&-\sum_{i=1}^{k} \frac{\lambda(t) \exp\{-a_i\lambda(t)\}}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N} (a_i - a_\alpha)} \\ &+ \sum_{i=1}^{k} \sum_{\substack{j=k+1 \\ a_i\neq a_j}}^{N} \frac{\exp\{-a_i\lambda(t)\} - \exp\{-a_j\lambda(t)\}}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha) \prod_{\delta=k+1,\delta\neq j}^{N} (a_j - a_\delta)(a_i - a_j)} \Bigg] \beta(t)\,dt, \end{aligned} \tag{40}$$

where $k = N/2$, and

$$\begin{aligned} F_{N+1}(t) = (-1)a_1 \cdots a_N \Bigg[&\sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N} (a_i - a_\alpha)} \left\{ \frac{\lambda(t)}{a_i} \exp\{-a_i\lambda(t)\} - \frac{1 - \exp\{-a_i\lambda(t)\}}{a_i^2} \right\} \\ &+ \sum_{i=1}^{k} \sum_{\substack{j=k+1 \\ a_i\neq a_j}}^{N} \frac{1}{(a_i - a_j) \prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha) \prod_{\delta=k+1,\delta\neq j}^{N} (a_j - a_\delta)} \left\{ \frac{1 - \exp\{-a_i\lambda(t)\}}{a_i} - \frac{1 - \exp\{-a_j\lambda(t)\}}{a_j} \right\} \Bigg] \end{aligned} \tag{41}$$

for $0 < t < \infty$. As $t \to \infty$, $f_{N+1}(t) \to 0$ and $F_{N+1}(t) \to 1$. To prove the last assertion, we take the limit of (41) as $t \to \infty$,

$$\begin{aligned} F_{N+1}(\infty) = (-1)a_1 \cdots a_N \Bigg[&\sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N} (a_i - a_\alpha)} \left(-\frac{1}{a_i^2}\right) \\ &- \sum_{i=1}^{k} \sum_{\substack{j=k+1 \\ a_i\neq a_j}}^{N} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha) a_i \prod_{\delta=k+1,\delta\neq j}^{N} (a_j - a_\delta) a_j} \Bigg]. \end{aligned} \tag{42}$$

Since $a_j = a_{N+1-j}$ and $k = N/2$, the limits of j (in the summation) and δ (in the product) in (42) may be changed from $(k + 1, N)$ to $(k, 1)$, and (42) may be rewritten

$$F_{N+1}(\infty) = a_1 \cdots a_N \left[\sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha)a_i} \right]^2, \tag{43}$$

where $k = N/2$,

$$a_1 \cdots a_N = (a_1 \cdots a_k)^2, \tag{44}$$

and, in light of Lemma 1,

$$\sum_{i=1}^{k} \frac{1}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha)a_i} = \frac{-1}{\prod_{i=1}^{k} (-a_i)}. \tag{45}$$

Substituting (44) and (45) in (43) yields

$$F_{N+1}(\infty) = 1. \tag{46}$$

In the same manner, it can be shown that whatever may be $n = 2, \cdots, N + 1$, $f_n(t) \to 0$ and $F_n(t) \to 1$ as $t \to \infty$, and the corresponding random variables T_n are all proper.

The expectation and variance of T_n can be computed directly from

$$E(T_n) = \int_0^\infty t f_n(t)\, dt \tag{47}$$

and

$$\sigma_{T_n}^2 = \int_0^\infty [t - E(T_n)]^2 f_n(t)\, dt. \tag{48}$$

For the duration of epidemic T_{N+1} with $N = 2k$, for example, we substitute (40) in (47) to obtain the expectation

$$E(T_{N+1}) = (-1)a_1 \cdots a_N \left[- \sum_{i=1}^{k} \frac{\int_0^\infty t\lambda(t) \exp\{-a_i\lambda(t)\}\beta(t)\, dt}{\prod_{\alpha=1,a_\alpha\neq a_i}^{N} (a_i - a_\alpha)} + \sum_{i=1}^{k} \sum_{j=k+1}^{N} \frac{\int_0^\infty t(\exp\{-a_i\lambda(t)\} - \exp\{-a_j\lambda(t)\})\beta(t)\, dt}{\prod_{\alpha=1,\alpha\neq i}^{k} (a_i - a_\alpha) \prod_{\delta=k+1,\delta\neq j}^{N} (a_j - a_\delta)(a_i - a_j)} \right]. \tag{49}$$

Obviously, explicit formulas of $E(T_n)$ and $\sigma_{T_n}^2$ depend upon the infection rate $\beta(t)$. When the infection rate is independent of time so that $\beta(t) = \beta$, the corresponding formulas may be obtained by an alternative method.

The length of time elapsed till the occurrence of the nth infection may be divided into two periods: a period of length T_{n-1} up to the occurrence of the $(n - 1)$th infection and a period of length W_n between the occurrence of the $(n - 1)$th and the nth infections. The sum of the two periods is equal to the entire length of time, or

$$T_n = T_{n-1} + W_n. \tag{50}$$

Equality (50) can be easily verified. When $\beta(t) = \beta$, T_{n-1} and W_n are independently distributed nonnegative random variables. The density functions of T_{n-1} and W_n can be derived from (34); they are

$$f_{n-1}(t) = P_{1,n-2}(0, t)a_{n-2}\beta \tag{51}$$

and

$$g_n(t) = P_{n-1,n-1}(0, t)a_{n-1}\beta, \tag{52}$$

respectively. According to (50), the distribution of T_n is the convolution of the distributions of T_{n-1} and W_n. Therefore, the corresponding density functions satisfy the relationship

$$f_n(t) = \int_0^t f_{n-1}(\tau)g_n(t - \tau)\,d\tau. \tag{53}$$

To prove (53), we recall identity (7) in Lemma 2,

$$P_{1,n-1}(0, t) = \int_0^t P_{1,n-2}(0, \tau)a_{n-2}\beta P_{n-1,n-1}(\tau, t)\,d\tau, \tag{54}$$

and multiply both sides of (54) by $a_{n-1}\beta$ to obtain

$$P_{1,n-1}(0, t)a_{n-1}\beta = \int_0^t [P_{1,n-2}(0, \tau)a_{n-2}\beta][P_{n-1,n-1}(\tau, t)a_{n-1}\beta]\,d\tau, \tag{55}$$

which, in light of (34), (51), and (52), is identical to (53), proving (50). Equation (50) is a special case of a general equality, for which the reader is referred to [6], p. 110.

Now, the probability in (52) is

$$P_{n-1,n-1}(0, t) = \exp\{-a_{n-1}\beta t\}; \tag{56}$$

therefore, the random variable W_n has an exponential distribution with the density function

$$g_n(t) = a_{n-1}\beta \exp\{-a_{n-1}\beta t\}. \tag{57}$$

The expectation and the variance of W_n, thus, are given by

$$E(W_n) = \frac{1}{a_{n-1}\beta} \tag{58}$$

and

$$\sigma^2_{W_n} = \frac{1}{a^2_{n-1}\beta^2}, \tag{59}$$

respectively.

Equation (50) can be easily extended. Let

$$W_i = T_i - T_{i-1}, \qquad i = 2, \cdots, N + 1, \tag{60}$$

be the length of time elapsed between the $(i - 1)$th and the ith infections. Using the arguments in proving (50), we can show that

$$T_n = W_2 + \cdots + W_n, \qquad n = 2, \cdots, N + 1, \tag{61}$$

where the W_i are independently distributed random variables, and each has an exponential distribution (see equation (57)) with

$$E(W_i) = \frac{1}{a_{i-1}\beta}, \qquad \sigma^2_{W_i} = \frac{1}{a^2_{i-1}\beta^2}, \qquad i = 2, \cdots, N+1. \tag{62}$$

It follows that the expectation and variance of T_n are

$$E(T_n) = \sum_{i=1}^{n-1} \frac{1}{a_i\beta}, \qquad \sigma^2_{T_n} = \sum_{i=1}^{n-1} \frac{1}{a_i^2\beta^2}, \quad n = 2, \cdots, N+1. \tag{63}$$

For the duration of the epidemic T_{N+1}, we may use the relationship $a_i = a_{N+1-i}$ to have

$$E(T_{N+1}) = 2\sum_{i=1}^{k} \frac{1}{a_i\beta}, \qquad \sigma^2_{T_{N+1}} = 2\sum_{i=1}^{k} \frac{1}{a_i^2\beta^2}, \tag{64}$$

when N is even with $k = N/2$, and

$$E(T_{N+1}) = 2\sum_{i=1}^{k-1} \frac{1}{a_i\beta} + \frac{1}{a_k\beta}, \qquad \sigma^2_{T_{N+1}} = 2\sum_{i=1}^{k-1} \frac{1}{a_i^2\beta^2} + \frac{1}{a_k^2\beta^2}, \tag{65}$$

when N is odd with $k = (N+1)/2$. They are the same as those derived from the cumulant generating function in [2], p. 47.

REFERENCES

[1] N. T. J. Bailey, "A simple stochastic epidemic," *Biometrika*, Vol. 37 (1950), pp. 193–202.
[2] ———, *The Mathematical Theory of Epidemics*, London, Charles Griffin, 1957.
[3] ———, "The simple stochastic epidemic: a complete solution in terms of known functions," *Biometrika*, Vol. 50 (1963), pp. 235–240.
[4] M. S. Bartlett, "Deterministic and stochastic models of recurrent epidemics," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 4, pp. 81–109.
[5] C. L. Chiang, "A stochastic model of competing risks of illness and competing risks of death," *Stochastic Models in Medicine and Biology* (edited by J. Gurland), Madison, University of Wisconsin Press, 1964, pp. 323–354.
[6] ———, *Introduction to Stochastic Processes in Biostatistics*, New York, Wiley, 1968.
[7] ———, "An equality in stochastic processes," *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1972, Vol. 4, pp. 187–196.
[8] H. E. Daniels, "The distribution of the total size of an epidemic," *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1965, Vol. 4, pp. 281–293.
[9] F. Downton, "A note on the ultimate size of a general stochastic epidemic," *Biometrika*, Vol. 54 (1967), pp. 314–316.
[10] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 2, New York, Wiley, 1965.
[11] J. Gani, "On a partial differential equation of epidemic theory," *Biometrika*, Vol. 52 (1965), pp. 617–620.
[12] D. G. Kendall, "Deterministic and stochastic epidemics in a closed population," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 4, pp. 149–165.
[13] J. Neyman and E. L. Scott, "A stochastic model of epidemics," *Stochastic Models in*

Medicine and Biology (edited by J. Gurland), Madison, University of Wisconsin Press, 1964, pp. 45–83.
[14] G. PÓLYA and G. SZEGÖ, *Aufgaben und Lehrsätze aus der Analysis*, Vol. 2, Berlin, Springer-Verlag, 1964.
[15] V. SISKIND, "A solution of the general stochastic epidemic," *Biometrika*, Vol. 52 (1965), pp. 613–616.
[16] P. WHITTLE, "The outcome of a stochastic epidemic," *Biometrika*, Vol. 42 (1955), pp. 116–122.
[17] G. L. YANG, "On the probability distributions of some stochastic epidemic models," TR 70-141, Department of Mathematics, University of Maryland, 1970.