An epidemic model approximating the spread of the common cold

J. Gani* and R.J. Swift**

Abstract

An approximation to the simple stochastic epidemic is presented which leads to a binomial type process for the number of susceptibles $X(t)$ at time $t \geq 0$. The expectation of $X(t)$ is compared with that of the exact simple epidemic, and the duration T of the epidemic is discussed.

Introduction

A simplified model for the spread of the common cold is the simple epidemic described by Bailey [2] and Daley and Gani [3]. This assumes the homogeneous mixing of cold-free susceptibles $X(t)$, where $X(0) = N$, and infectives with the cold $Y(t)$, where $Y(0) = I$. As examples, we may consider a family of six members setting off the epidemic with $N = 5$ and $I = 1$, or alternatively, a classroom of 21 members, starting with $N = 20$, $I = 1$. In the deterministic version of the model, which is characterised by the differential equations

$$\frac{\mathrm{d}X}{\mathrm{d}t} = -\beta XY\,, \qquad \frac{\mathrm{d}Y}{\mathrm{d}t} = \beta XY\,,$$

it is well known that for the infection parameter $\beta = 1$,

$$X(t) = \frac{N(N+I)}{N + I\mathrm{e}^{(N+I)t}}, \qquad Y(t) = \frac{I(N+I)\mathrm{e}^{(N+I)t}}{N + I\mathrm{e}^{(N+I)t}}\,, \tag{1}$$

with an epidemic curve, representing the rate of spread of the infection,

$$w = XY = -\frac{\mathrm{d}X(t)}{\mathrm{d}t} = \frac{NI(N+I)^2\mathrm{e}^{(N+I)t}}{(N + I\mathrm{e}^{(N+I)t})^2}\,. \tag{2}$$

Bailey [2, p. 35] provided graphs of w when $N = 10$, $I = 1$ and $N = 20$, $I = 1$; these graphs are redrawn in Figure 1.

Received 8 August 2007; accepted for publication 7 November 2007.
*Mathematical Sciences Institute, Australian National University, Canberra ACT 0200.
E-mail: gani@maths.anu.edu.au
**Department of Mathematics and Statistics, California State Polytechnic University, Pomona, Pomona, CA 91768, USA. E-mail: rjswift@csupomona.edu

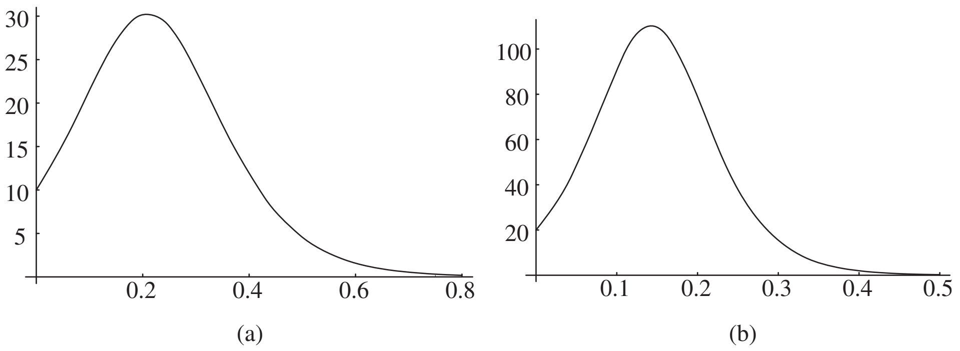

Figure 1. The deterministic epidemic curve for two initial susceptible population sizes: (a) $N = 10$, $I = 1$; (b) $N = 20$, $I = 1$.

The simple stochastic epidemic

In a stochastic formulation of the simple epidemic, $X(t)$ and $Y(t)$ are considered to be random variables (RVs), where the probabilities

$$\mathrm{P}_i(t) = \mathrm{P}\{X(t) = i \mid X(0) = N\}$$

of the Markov chain $\{X(t); t \geq 0\}$ in continuous time will satisfy the equations

$$\frac{\mathrm{d}P_i(t)}{\mathrm{d}t} = -i(N + I - i)\,\mathrm{P}_i(t) + (i+1)(N + I - i - 1)\,\mathrm{P}_{i+1}(t)\,, \quad 0 \leq i \leq N-1\,,$$
$$\frac{\mathrm{d}P_N(t)}{\mathrm{d}t} = -iN\,\mathrm{P}_N(t)\,. \tag{3}$$

Writing the probability generating function (PGF) of $X(t)$ as

$$\phi(z,t) = \sum_{i=0}^{N} \mathrm{P}_i(t) z^i\,, \quad 0 \leq z \leq 1$$

with $\phi(z,0) = z^N$, we can derive the partial differential equation (PDE)

$$\frac{\partial \phi}{\partial t} = z(z-1)\frac{\partial^2 \phi}{\partial z^2} - (z-1)(N+I-1)\frac{\partial \phi}{\partial z} \tag{4}$$

which yields on differentiation with respect to z, when $z = 1$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{E}[X(t)] = -\,\mathrm{E}[X(t)](N + I - \mathrm{E}[X(t)]) + \mathrm{var}[X(t)] \tag{5}$$

as against the deterministic

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = -XY = -X(N+I-X)\,.$$

Bailey [2] derived the PGF ϕ, the exact solution of equation (4), as well as the moment generating function (MGF) of $X(t)$ in terms of hypergeometric functions. He also obtained the approximate solution using a perturbation technique that made a small change in the states of the process by letting $n = N + \epsilon$ where $\epsilon > 0$. This gives an approximate value for the PGF ϕ as

$$\phi(z,t) = \sum_{j=0}^{N} \mathrm{d}_j \mathrm{e}^{-j(n+1-j)t}\,{}_2F_1(-j, j-n-1, -n, z)\,, \tag{6}$$

where

$$\mathrm{d}_j = \frac{(-1)^j N!\,(n-2j+1)n!}{j!(N-j)!\,(n-N)!\,\prod_{r=0}^{N}(n-j-r+1)}\,,$$

and ${}_2F_1(-j, j-n-1, -n, z)$ is the hypergeometric function defined by the series

$${}_2F_1(a,b,c,z) = \sum_{\ell=0}^{\infty} \frac{(a)_\ell\,(b)_k}{(c)_\ell}\,\frac{z^\ell}{\ell!}$$

with $(a)_\ell = a(a+1)(a+2)\cdots(a+\ell-1)$ and $(a)_0 = 1$.

Bailey [2] and Daley and Gani [3] provide details of this solution and both show that the approximation tends to the exact solution as ε goes to zero.

The expected number of susceptibles for the simple stochastic epidemic is obtained from (6) as

$$\mathrm{E}[X(t)] = \frac{\partial\phi(1,t)}{\partial z} = \sum_{j=0}^{N} \mathrm{d}_j \mathrm{e}^{-j(n-j+1)t} \sum_{k=1}^{\infty} \frac{(-j)_k(j-n-1)_k}{(-n)_k(k-1)!}\,.$$

The epidemic curve for the simple stochastic epidemic process is found as

$$w = -\frac{\mathrm{d}\,\mathrm{E}[X(t)]}{\mathrm{d}t}$$

which is graphed in Figure 2 for $N = 10$, $I = 1$ and $N = 20$, $I = 1$, with $\varepsilon = 0.001$.

In the next section, we propose a much simpler alternative approximation in which the RV $Y(t)$ is replaced by its deterministic value. This will provide a more readily tractable model for the epidemic.

An approximate simple stochastic epidemic

In the previous section, (5) shows that if $\mathrm{var}[X(t)]$ is small compared to the product $X(N+I-X)$, then the deterministic model for the number of susceptibles and the expected number of susceptibles for the stochastic model will be close. Thus, we use the idea of replacing an RV by its deterministic value as the basis of an approximation to the simple stochastic epidemic.

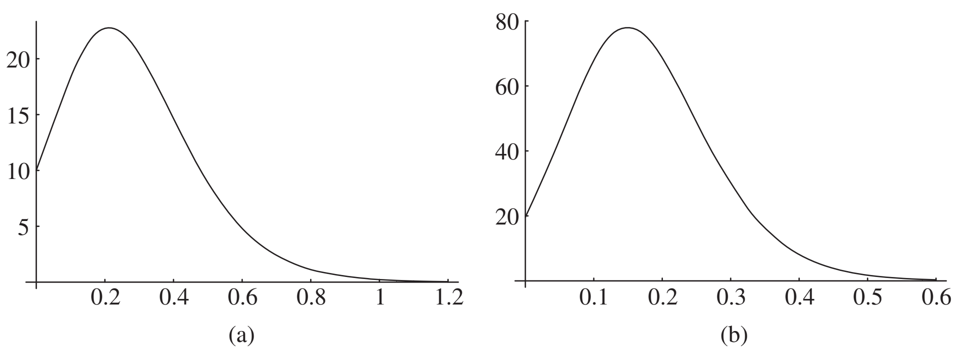

Figure 2. The stochastic epidemic curve for two initial susceptible population sizes: (a) $N = 10$, $I = 1$; (b) $N = 20$, $I = 1$. In both graphs $\varepsilon = 0.001$.

Specifically, for the transition of the number of susceptibles from i to $i-1$, we replace the probability $i(N+I-i)\delta t + o(\delta t)$ by the probability

$$i\left(\frac{(N+I)I\mathrm{e}^{(N+I)t}}{N+I\mathrm{e}^{(N+I)t}}\right)\delta t + o(\delta t)\,.$$

This transforms the process $X(t)$ into a pure death process with a time-dependent death parameter

$$\mu(t) = \frac{I(N+I)\mathrm{e}^{(N+I)t}}{N+I\mathrm{e}^{(N+I)t}} \tag{7}$$

and binomial type PGF

$$F(z,t) = \left((z-1)\exp\left(-\int_0^t \mu(v)\,\mathrm{d}v\right)+1\right)^N = \left(\frac{(z-1)(N+I)}{N+I\mathrm{e}^{(N+I)t}}+1\right)^N, \tag{8}$$

where

$$\int_0^t \mu(v)\,\mathrm{d}v = \ln\left(\frac{N+I\mathrm{e}^{(N+I)t}}{N+I}\right) \quad \text{and} \quad \exp\left(\int_0^t \mu(v)\,\mathrm{d}v\right) = \frac{N+I\mathrm{e}^{(N+I)t}}{N+I}\,. \tag{9}$$

Further details on birth–death processes with time-dependent parameters are given by Allen [1].

Note that the expectation $\mathrm{E}[X(t)]$ of this approximate process takes exactly the deterministic value in (1), so that the epidemic curve for this approximate simple stochastic epidemic is precisely that given by (2).

We now compare this expectation with the exact expectation derived by Bailey for $N=10$, $I=1$ and $N=20$, $I=1$. These graphs are presented in Figure 3. Note that the approximate expectations of the numbers of susceptibles are underestimates of the exact values by just over 1 when $N=10$, and about 2 when $N=20$.

Duration of the epidemic

The simple stochastic epidemic process evolves by unit decrements at the times $t_N, t_{N-1}, \ldots, t_1$ with $t_{N+1}=0$, as shown in Figure 4. The duration time T of the

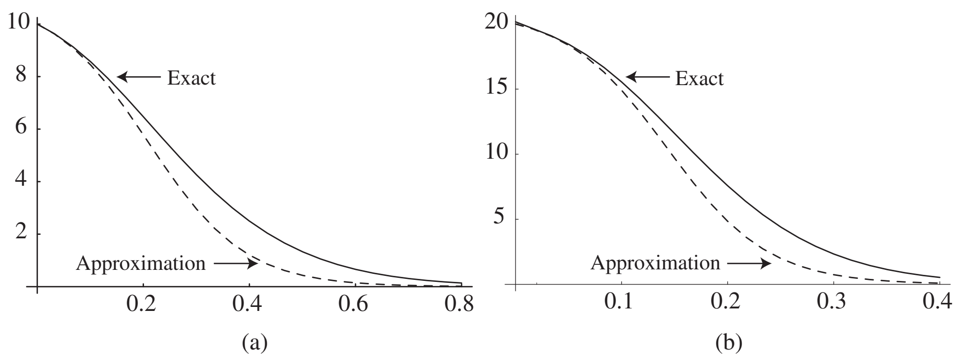

Figure 3. The expected value $E[X(t)]$ of the exact stochastic simple epidemic is compared with the equivalent expected value for the approximate stochastic simple epidemic for two initial population sizes of susceptibles: (a) $N=10$, $I=1$; (b) $N=20$, $I=1$.

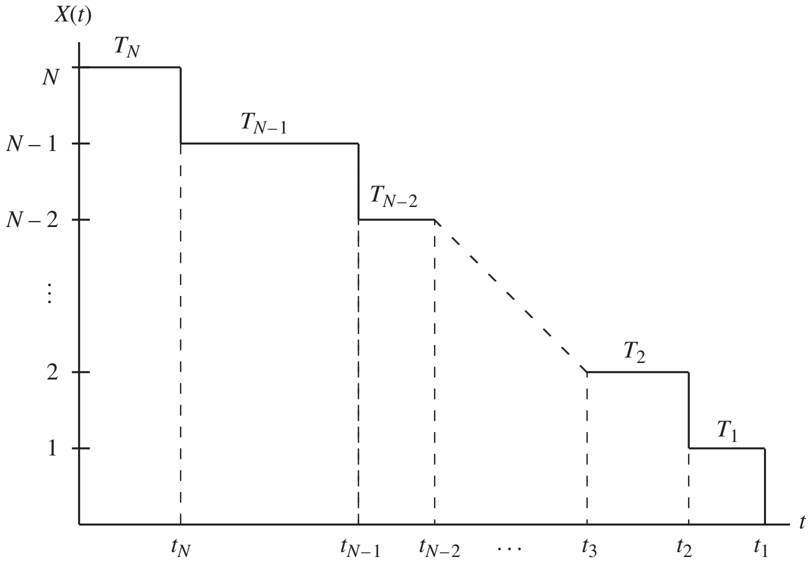

Figure 4. The simple stochastic epidemic process.

epidemic, is given by $T = \sum_{i=1}^{N} T_i$ where the $T_i = t_i - t_{i+1}$, $(i = 1, 2, \ldots, N)$ are independent exponentially distributed random variables, with

$$\mathrm{E}[T_i] = \frac{1}{i(N + I - i)}\,,$$

so that

$$\mathrm{E}[T] = \sum_{i=1}^{N} \frac{1}{i(N + I - i)}. \tag{10}$$

To find the duration $\tilde{T}$ of the approximate simple stochastic epidemic, we write the PGF $F(z,t)$ in (8) as

$$F_X(z,t) = \left(\frac{z(N + I) + I(\mathrm{e}^{(N+I)t} - 1)}{N + I\mathrm{e}^{(N+I)t}} \right)^N$$

which gives that

$$\mathrm{P}_1(t) = \mathrm{P}\{X(t) = 1 \mid X(0) = N\} = \frac{N(N + I)I^{N-1}(\mathrm{e}^{(N+I)t} - 1)^{N-1}}{(N + I\mathrm{e}^{(N+I)t})^N}\,.$$

The probability density function of the RV $\tilde{T}$ for the approximate simple stochastic epidemic is given by

$$f(t) = \mathrm{P}_1(t)(N + I - 1) = \left(\frac{NI^{N-1}(N + I)(N + I - 1)(1 - \mathrm{e}^{-(N+I)t})^{N-1}}{\mathrm{e}^{(N+I)t}(I + N\mathrm{e}^{-(N+I)t})^N} \right),$$

so that the expected duration is found to be

$$\begin{aligned}
\mathrm{E}[\tilde{T}] &= \int_0^\infty t f(t)\,\mathrm{d}t \\
&= \int_0^\infty t\left(\frac{NI^{N-1}(N+I)(N+I-1)(1-\mathrm{e}^{-(N+I)t})^{N-1}}{\mathrm{e}^{(N+I)t}(I+N\mathrm{e}^{-(N+I)t})^N}\right)\mathrm{d}t \\
&= \left(\frac{N}{I}\right)(N+I)(N+I-1)\sum_{j=0}^{N-1}\binom{N-1}{j}(-1)^j \\
&\quad\times\int_0^\infty t\mathrm{e}^{-(N+I)(j+1)t}\left(1-\left(\frac{N}{I}\right)\mathrm{e}^{-(N+I)t}\right)^{-N}\mathrm{d}t \\
&= \left(\frac{N(N+I-1)}{I(N+I)}\right)\sum_{j=0}^{N-1}\sum_{k=0}^{\infty}\binom{N-1}{j}\binom{N+k}{k}(-1)^{j+k}\left(\frac{I}{N}\right)^k\frac{1}{(j+k+1)^2}\,.
\end{aligned} \tag{11}$$

While our model leads to simpler transient probabilities, the expected duration given in (11) has a more complicated structure than that shown in (10).

Table 1 compares this approximation for the expected duration with the exact expected duration from (10), we see that for large N, the approximate duration is very close to the exact value.

Table 1. Expected duration times of the exact and approximate simple stochastic epidemic process with $I = 1$.

Initial bumber of susceptibles	$N = 10$	$N = 20$	$N = 100$	$N = 500$
Exact duration	0.5325	0.3426	0.10272	0.02711
Approximate duration	0.4667	0.3116	0.0994	0.02636

It should be pointed out that some asymptotic results are available for the distribution of T; for $I = 1$, Bailey [2] points out that the variable

$$W = (N+1)T - 2\ln N$$

has the approximate distribution function

$$H(w) = 2\mathrm{e}^{-\frac{1}{2}w}K_1(2\mathrm{e}^{-\frac{1}{2}w})\,,$$

where K_1 is the modified Bessel function of the second kind. (See [4] for the original formulae for the distribution of W.)

References

[1] Allen, L.J.S. (2003). *An Introduction to Stochastic Processes with Applications to Biology.* Pearson Education, Upper Saddle River, NJ.
[2] Bailey, N.T.J. (1975). *The Mathematical Theory of Infectious Disease.* Griffin, London.
[3] Daley, D.J. and Gani, J. (1999). *Epidemic Modelling.* Cambridge University Press.
[4] Kendall, D.G. (1957). La propagation d'une épidémie ou d'un bruit dans une population limitée. *Publ. Inst. Statist. Univ. Paris* **6,** 307–311.