DETERMINISTIC AND STOCHASTIC EPIDEMICS IN CLOSED POPULATIONS

DAVID G. KENDALL
MAGDALEN COLLEGE, OXFORD UNIVERSITY

1. Introduction

The problem of the growth of a stochastic epidemic in a closed population is a very challenging one; from one point of view it is almost trivial, for we have only to deal with a temporally homogeneous Markov process having a *finite* number of states, and yet there is great difficulty in finding out anything useful about the sample "epidemic curve" (the plot of the number of infectious individuals in circulation in the population as a function of the time[1]). We are much better informed about the ultimate behaviour of the system; the ordinate of the epidemic curve must eventually fall to zero and then the system will be "frozen" with say $X(\infty)$ "susceptibles" and $Z(\infty)$ "removed" persons, where $X(\infty) + Z(\infty)$ is equal to the (fixed) population size, and the form of the distribution of the random variable $Z(\infty)$ has been investigated by McKendrick [9] and by Bailey [1], [2]; the latter has even given an explicit formula for this distribution (due to F. G. Foster). So we know the ultimate fate of such a system, but little about how that state is attained. The difficulties blocking the way to an analytical solution are made clear by Professor Bartlett in his contribution to the present Symposium. In this paper I wish to sketch some approximate procedures which give a partial answer, which may be of value in more complicated situations of the same sort and which it would be interesting to replace by a mathematical argument. It is to be hoped that no one will apply them without reflexion, for there is nothing to be said in their favour beyond a certain intuitive plausibility and the questionable confirmation afforded by an armchair Monte Carlo experiment.

2. Deterministic epidemics

We must begin by making a thorough study of the corresponding *deterministic* system; this is the closed epidemic in the form in which it was studied by Kermack and McKendrick [8]. These writers partitioned the population of n individuals into three classes:

x susceptibles,

y infected persons in circulation in the population, and

z "removed" persons (recovered, dead or isolated),

so that

$$x + y + z = n = \text{constant}. \tag{1}$$

[1] This definition of "epidemic curve" and the definition of the "notifications curve" in section 2 may differ from the reader's own usage.

It is supposed that there is migration from the x-class to the y-class (*infection*) at a rate proportional to xy, and removal from the y-class to the z-class (*recovery*, etc.) at a rate proportional to y; there is no exit from the z-class and no entry into the x-class. Thus we have the Kermack-McKendrick (K and K) equations,

$$\frac{dx}{dt} = -\beta xy \ ,$$

$$\frac{dy}{dt} = \beta xy - \gamma y \ , \tag{2}$$

$$\frac{dz}{dt} = \gamma y \ ,$$

where β and γ are constants, and immediately it follows that

$$x = x_0 e^{-(z-z_0)/\rho}, \tag{3}$$

where $\rho \equiv \gamma/\beta$. If initially $x = x_0$, $y = y_0$ and $z_0 = 0$ then

$$\frac{dz}{dt} = \gamma(n - z - x_0 e^{-z/\rho}) \ . \tag{4}$$

K and K dealt with this equation by replacing the exponential by the first three terms of its Taylor series; the resulting equation can be integrated explicitly to give, as an approximation to the epidemic curve,

$$y = C \operatorname{sech}^2(\alpha t - \phi) \tag{5}$$

where C, α and ϕ are constants. It has often been noted that in practice we observe not $y(t)$ but $(d/dt)z(t)$ (the "*notifications*" curve[2]); however the ordinates of the two curves will be in a constant proportion in virtue of the third of equations (2) (although they will *not* be so related in the stochastic case). It is curious that the K and K approximation (5) should have been accepted without comment for nearly thirty years; the exact solution is easily obtained and the difference between the two can be of practical significance.

An indirect way of examining the appropriateness of an approximation is to determine the modified model to which it is the exact solution. To this end let us temporarily assume that the infection-rate β varies systematically with z (this is almost equivalent to a dependence of the infection-rate on the "age" of the epidemic). In place of (3) and (4) we shall now have

$$\log(x/x_0) = -\frac{1}{\gamma}\int_0^z \beta(w)dw \tag{6}$$

and

$$\frac{dz}{dt} = \gamma\left(n - z - x_0 \exp\left[-\frac{1}{\gamma}\int_0^z \beta(w)dw\right]\right) . \tag{7}$$

The differential equation (7) will coincide with the K and K approximation to (4) when

$$\beta(z) \equiv \frac{2\beta}{[1 - z/\rho] + [1 - z/\rho]^{-1}} \ , \tag{8}$$

[2] See footnote 1.

and so $\beta(z) = \beta$ when $z = 0$ and $\beta(z) < \beta$ when $0 < z \leqq \rho$. From this we infer that the K and K approximation consistently *under*estimates the infection-rate and so it will underestimate the total size $z(\infty)$ of the epidemic, as may also be seen more directly from (4). It will correspond to a hopelessly unrealistic model if at any time $z > \rho$ (for then a negative infection-rate will be in operation), and the infection-rate will be kept within 10 per cent of its initial value only if $z(\infty) \leqq 0.373\ \rho$. Now the K and K approximation leads to the formula

$$\lim_{y_0 \to 0} z(\infty) = 2\rho(x_0 - \rho)/x_0 \qquad (x_0 > \rho)\ , \tag{9}$$

$$= 0 \qquad (x_0 \leqq \rho)\ ,$$

and so, accepting this as a rough relation between x_0 and $z(\infty)$, we cannot expect a consistent approximation if $x_0 > 2\rho$, and we must be prepared for appreciable errors in the infection-rate if $x_0 > 1.23\rho$. In previous work, values of x_0/ρ up to 4 have been considered; this makes evident the need for an accurate treatment of equations (2), and we shall now give one.

3. The exact solution to the Kermack-McKendrick equations

It is usual to discuss the epidemic generated by the addition of a given (small) number of infectious persons to a population consisting entirely of susceptibles. However it is a feature of the deterministic model that some of the formulas give meaningful results even in the limit when the initial number of infectious persons (y_0) is allowed to tend to zero; thus it is possible to predict the size of epidemic (that is, the total number of "cases") resulting from the introduction of a mere trace of infection into a wholly susceptible population of size x_0 when the constant ρ is given. This is, in fact, what the K and K formula (9) does, for their approximating model. We have seen that (9) must not be used if $x_0 \gg \rho$ and in practice it is most frequently of interest when $x_0 \cong \rho$; in this case a further approximation is possible and we can write

$$\lim_{y_0 \to 0} z(\infty) = 2(x_0 - \rho) \qquad (x_0 > \rho)\ , \tag{9a}$$

$$= 0 \qquad (x_0 \leqq \rho)\ .$$

Equation (9a) asserts (i) that there will be no epidemic if x_0 is less than ρ (the so-called *threshold*) and (ii) that in all other cases the number of susceptibles will fall as far below the threshold as it was initially above it. These two statements constitute the celebrated *threshold theorem* of K and K. One of our tasks will be to examine the exact behaviour of the deterministic model in these respects, and we shall then have to enquire what sort of threshold phenomena may be expected in the stochastic situation.

Let ζ_2 be the (unique) positive root of the equation

$$z = n - x_0 e^{-z/\rho}\ ; \tag{10}$$

then we shall have, on integrating (4),

$$\gamma t = \int_0^z \frac{dw}{n - w - x_0 e^{-w/\rho}} \qquad (0 \leqq z < \zeta_2)\ , \tag{11}$$

and this when taken in conjunction with (4) gives a pair of parametric equations for the notifications curve relating dz/dt to t. The whole curve for $0 \leqq t < \infty$ is so represented because the integral in (11) diverges when $z \to \zeta_2$; ζ_2 is accordingly the correct value of $z(\infty)$. If now we put $y_0 = 0$ (so that $n = x_0$) in (4) and (11), the integral also diverges at its lower limit and we are in trouble because the epidemic takes an infinitely long time to get started. We can avoid this difficulty by a change of time-origin. K and K in 1927 noted that

$$\frac{d}{dt}\left(\frac{dz}{dt}\right) = \gamma^2 y\left(\frac{x}{\rho} - 1\right), \tag{12}$$

so that *the peak of the notifications (and so also of the epidemic) curve occurs when the number of susceptibles is equal to the threshold* (ρ). Let us call this epoch the *centre* of the epidemic and adopt it as the origin of time. We shall now have $x_0 = \rho$; the value of y_0 will be an important parameter distinguishing between epidemics of varying severity (it is, with the natural time-unit $1/\gamma$, the peak height of the notifications curve) and without any loss of generality it will still be convenient to take $z_0 = 0$. The *numerical* value of $z(t)$ will then be the number of removed cases in $(0, t)$ if t is positive and in $(t, 0)$ if t is negative. The notifications curve will now be given parametrically by the pair of equations

$$\begin{aligned} \gamma t &= \int_0^z \frac{dw}{y_0 + \rho(1 - e^{-w/\rho}) - w}, \\ \frac{1}{\gamma}\frac{dz}{dt} &= y_0 + \rho(1 - e^{-z/\rho}) - z, \end{aligned} \tag{13}$$

where $-\infty < t < \infty$ and $-\zeta_1 < z < \zeta_2$, and $-\zeta_1$ and ζ_2 are the (unique) negative and positive roots of the equation

$$e^{-\zeta/\rho} - 1 + \zeta/\rho = y_0/\rho . \tag{14}$$

The complete change in our point of view should be noted; to begin with we followed the customary procedure of studying the development of an epidemic subsequent to its artificial creation by the introduction of y_0 infectious persons into a population of x_0 susceptibles. We are now considering an epidemic as an entity existing from $t = -\infty$ to $t = +\infty$, and we have for convenience located our time-origin at the epoch of greatest activity. When we come to examine the analogous stochastic system we shall return to the earlier point of view, and examine epidemics started up artificially at $t = 0$.

It will be seen from (13) that

$$\lim_{t \to -\infty} z(t) = -\zeta_1, \quad \lim_{t \to \infty} z(t) = \zeta_2, \tag{15}$$

and so ζ_1 and ζ_2 are the numbers of cases removed from circulation *before* and *after* the central epoch, respectively. The asymmetry of the notifications curve will be betrayed by an inequality between ζ_1 and ζ_2 (we shall find that $\zeta_1 < \zeta_2$), and this is the first point at which we notice a qualitative departure from the K and K solution (5).

The approximate shape of the notifications curve near its peak (at $t = 0$) can be found by applying a K and K type of approximation to (13); this gives

$$\frac{1}{\gamma}\frac{dz}{dt} \cong y_0 \operatorname{sech}^2\left\{\gamma t \sqrt{\frac{y_0}{2\rho}}\right\}, \tag{16}$$

an approximation valid so long as $|z|/\rho$ is small.

TABLE I

DETERMINISTIC EPIDEMICS OF THE K AND K TYPE

Proportion of Population Attacked I per cent	Population as Multiple of "Threshold" N/ρ	Proportion Infectious and in Circulation at "Central" Epoch y_{max}/N per cent	Proportion of Removals Occurring before "Central" Epoch $\zeta_1/(\zeta_1+\zeta_2)$ per cent
0	1.00	0.0	50
10	1.05	0.1	49
20	1.12	0.6	49
30	1.19	1.3	48
40	1.28	2.5	48
50	1.39	4.3	47
60	1.53	6.8	46
70	1.72	10.3	45
80	2.01	15.5	43
90	2.56	24.2	41
95	3.2	31.9	38
98	4.0	40.3	35

When $t = -\infty$, the number N of susceptibles must have been $\rho + y_0 + \zeta_1$, and this now plays the role of the total population size. Thus

$$I \equiv \frac{\zeta_1 + \zeta_2}{\rho + y_0 + \zeta_1} \tag{17}$$

is the fraction of the total population which succumbs to the disease (either before or after the central epoch) and we shall call I the *intensity* of the epidemic.

In table I will be found a few figures which illustrate the exact behaviour of a K and K epidemic. It will be found useful in performing such calculations to take the intensity I as a basic parameter. If I and ρ are given we consider first the transition:

Epoch	$t = -\infty$	$t = +\infty$
x-men	N	$N - NI$
y-men	0	0
z-men	0	NI

Equation (3) then gives at once

$$N - NI = Ne^{-NI/\rho} \tag{18}$$

so that

$$\frac{N}{\rho} = \frac{1}{I} \log \frac{1}{1 - I} . \tag{19}$$

We next consider the transition:

Epoch	$t = -\infty$	$t = 0$
x-men	N	ρ
y-men	0	y_0
z-men	0	ζ_1

and this time (3) gives

$$\rho = Ne^{-\zeta_1/\rho} . \tag{20}$$

Now $I = (\zeta_1 + \zeta_2)/N$, and so as a measure of the asymmetry we have

$$\frac{\zeta_1}{\zeta_1 + \zeta_2} = \frac{\rho}{NI} \log \frac{N}{\rho} . \tag{21}$$

Finally the number of infectious persons in circulation at the central epoch will be

$$y_0 = N - \rho - \zeta_1 , \tag{22}$$

and the peak height of the notifications curve will be

$$\left(\frac{dz}{dt}\right)_{\max} = \gamma(N - \rho - \zeta_1) . \tag{23}$$

These figures are very instructive and throw a good deal of light on the Kermack-McKendrick threshold theorem. The first thing to notice is that $N \geqq \rho$, so that the initial number of susceptibles in an epidemic generated at an infinitely remote epoch by the introduction of a trace of infection must always exceed the threshold. The number of susceptibles will fall to the threshold value when the peak occurs on the notifications curve, and it will then continue falling to a final value $N(1 - I)$. To take an example, let $\rho = 100$ and suppose that the population starts with $N = 105$ susceptibles and a trace of infection; in the resulting epidemic (which will take an infinitely long time to develop) 10 per cent of the population will succumb to the disease and so the final number of susceptibles will be 94.5. In this example, therefore, the crude K and K formula,

$$\rho - x(\infty) \cong x(-\infty) - \rho , \tag{24}$$

works very well. But if the intensity of the epidemic is large then the K and K rule fails. Thus, if $\rho = 100$ and if now the population starts at $t = -\infty$ with $N = 201$ susceptibles and a trace of infection, we see from table I that the epidemic will have an intensity $I = 80$ per cent; the final number of susceptibles will thus be 40. The crude K and K rule would have predicted zero as the final number of susceptibles and the more accurate approximation (9) would have predicted a figure of 100. For epidemics of considerable severity, therefore, neither of (9) and (9a) can safely be

used and the exact value should either be read off from table I or calculated by the method just explained.

The last column of table I provides information about the asymmetry of the notifications curve; this shows that when the initial number of susceptibles is as large as 4ρ then nearly two-thirds of the removals occur *after* the peak of the notifications curve. It would be of interest to know whether empirical notifications curves are skewed *in this direction.*

Table I is concerned only with epidemics extending from $t = -\infty$ to $t = +\infty$, but exactly the same principles can be used to find the subsequent behaviour of an epidemic which is artificially started at $t = 0$ by the addition of a infectious persons to a population of m susceptibles; a worked example will be found in the next section of this paper. *In these circumstances an epidemic will always be generated, whether or not m exceeds the threshold*; however, *the peak of the notifications curve will not be observed unless the induced epidemic is in a precentral phase.* This means that, for all values of a, the notifications curve will be peaked if $m > \rho$ and J-shaped if $m \leqq \rho$.

4. Stochastic epidemics

Let us turn now to the stochastic model, first considered by McKendrick [9] and amplified by Bartlett [3], [4] which is the natural stochastic equivalent to the deterministic scheme of sections 2 and 3. The numbers of susceptible persons, of infectious persons and of removed cases will now be integer-valued random variables to be denoted by $X(t)$, $Y(t)$ and $Z(t)$, respectively. The epidemic will be thought of as having been artificially started at the epoch $t = 0$ with the initial conditions

$$X(0) = m\,, \quad Y(0) = a\,, \quad Z(0) = 0\,,$$

and a distinction will be drawn between the *primary cases* introduced into the population from the outside and the *secondary cases*, $m - X(\infty)$ in number, which occur among the m original susceptibles. The stochastic process is to be a Markovian one with a finite number of states and its character will be sufficiently indicated by listing the transitions possible in the time interval $(t, t + dt)$:

(i) *an infection* ($X \to X - 1$ and $Y \to Y + 1$) with probability $\beta XY\,dt + o(dt)$;
(ii) *a removal* ($Y \to Y - 1$ and $Z \to Z + 1$) with probability $\gamma Y\,dt + o(dt)$;
(iii) a variety of *multiple transitions*, with total probability $o(dt)$; and
(iv) *no change*, with probability $1 - (\beta X + \gamma)Y\,dt + o(dt)$.

The ratio $\rho \equiv \gamma/\beta$ is again important and will still be called the *threshold.* (Bailey [2] calls it the relative removal rate.)

Bailey [1], [2] has investigated the distribution of the total number

$$w \equiv m - X(\infty) = Z(\infty) - a \tag{25}$$

of secondary cases and his calculations confirm and extend preliminary conclusions already reached by McKendrick [9]. Bailey's results were presented graphically in a form which proved difficult to describe in simple terms; the following reclassification of his results tells a clearer story. (Here and elsewhere, when referring to Mr. Bailey's work, I am indebted to him for permission to include some unpublished material.)

The J-shaped distributions have a peak at $w = 0$ and the ordinate then falls off steadily as w increases. The bimodal distributions have one peak at $w = 0$ and the ordinates then assume very small values until a second peak appears at or just before $w = m$. For $m/\rho = 4$ the two peaks are of comparable height and the distribution is U-shaped. Bailey's calculations thus suggest the occurrence of a definite change in form at $m = \rho$; it would be of interest to investigate this stochastic threshold phenomenon analytically, using the formula developed by Bailey and Foster, but this has not yet been done.[3] We can summarise Bailey's findings in another way and say that (i) there will be a minor epidemic if $m \leqq \rho$, and (ii) there will be *either* a minor *or* a major epidemic if $m > \rho$, epidemics of intermediate magnitude being very unlikely.

TABLE II

CHARACTER OF THE DISTRIBUTION OF THE TOTAL NUMBER, w, OF SECONDARY CASES
(Number, a, of primary cases = 1)

m	m/ρ					
	0.67	0.80	1	1.33	2	4
10 20 40	All J-shaped			All bimodal		

(Based on calculations by N. T. J. Bailey.)

A moment's reflexion shows that this mode of behaviour might well have been expected. At $t = 0$ the population of infectious individuals (Y-men) is subject to a "tangential" birth-and-death process with a constant death-rate equal to γ and a stochastic birth-rate equal to $\beta X(t)$. Thus, if m is reasonably large, we shall for a short time have in effect a birth-and-death process in which the ratio death-rate/birth-rate has the value

$$\mu/\lambda = \gamma/(\beta m) = \rho/m . \tag{26}$$

The chance of ultimate extinction in such a birth-and-death process is $(\rho/m)^a$ if $m > \rho$, and is unity if $m \leqq \rho$. Thus, if $m \leqq \rho$, we may expect a small number of secondary cases and then quiescence. If $m > \rho$, however, the tangential birth-and-death process will *either* die out quickly *or* grow indefinitely. Thus we may then expect with complementary probabilities that the number of Y-men will *either* fall speedily to zero, producing a small number of secondary cases, *or* build up to large values, so that a major epidemic results. If this view is correct, then the area under the first peak of the w-distribution should be approximately equal to $(\rho/m)^a$ when $m > \rho$. *Very roughly*, this is so; particulars will be found in table III. The agreement is quite reasonable for $m \geqq 2\rho$ and sufficiently so to make our idea seem worth pursuing.

Now when $m > \rho$ and when by chance the system displays its *second* mode of behaviour (large numbers of infectious persons being built up) then we can reasonably

[3] An important contribution to this problem has now been made by P. Whittle [10]; see, also, the paper by F. G. Foster [6].

approximate by assuming quasi-deterministic behaviour, and adapt the formulas of sections 2 and 3. If, however, the *first* mode of behaviour (almost immediate removal of the infectious persons) occurs then we can approximate by assuming that the number of Y-men will follow the tangential birth-and-death process *conditioned by the requirement of ultimate extinction.* In fact we propose the following stochastic system as an approximation to the one actually being studied:

TABLE III

THE CHANCE OF NOT MORE THAN ONE SECONDARY CASE, COMPARED WITH $(\rho/m)^a$ (SHOWN IN PARENTHESES)

($a = 1$)

m	m/ρ		
	1.33	2	4
10	0.55 (0.75)	0.42 (0.50)	0.24 (0.25)
20	0.54 (0.75)	0.41 (0.50)	0.23 (0.25)
40	0.54 (0.75)	0.41 (0.50)	0.23 (0.25)

(N. T. J. Bailey's calculations).

Specification of the approximating stochastic system:

(i) If $m \leqq \rho$, then $\tilde{Y}(t)$ is the number of individuals in a population controlled by a simple birth-and-death process with birth-rate $= m\beta$ and death-rate $= \gamma$, satisfying the initial condition $\tilde{Y}(0) = a$.

(ii) If $m > \rho$, then the system has two modes of behaviour, A and B, with

$$\text{pr}\,(A) = (\rho/m)^a \quad \text{and} \quad \text{pr}\,(B) = 1 - (\rho/m)^a. \tag{27}$$

In mode A, $\tilde{Y}(t)$ behaves as if it were the number of individuals in a simple birth-and-death process with birth-rate $= m\beta$ and death-rate $= \gamma$, satisfying the initial condition $\tilde{Y}(0) = a$ *and further conditioned by the requirement that* $\tilde{Y}(\infty) = 0$.

In mode B, $\tilde{Y}(t)$ behaves as if it were the function $y(t)$ describing the evolution of the corresponding deterministic epidemic.

Before discussing the behaviour of the approximating system it will be necessary to know how a birth-and-death process for which the birth-rate $\lambda(\equiv m\beta)$ exceeds the death-rate $\mu(\equiv\gamma)$ is affected when one is required (as in mode A above) to reject all sample-functions which do not eventually attain and thereafter remain constantly at the value zero. The solution to this problem has recently been found by W. A. O'N. Waugh [10]: *a birth-and-death process with a birth-rate* λ *which is greater than the death-rate* μ *coincides, when subject to the requirement of ultimate extinction, with an unconditioned birth-and-death process for which* μ *is now the birth-rate and* λ *the death-rate.* Thus *in the definition of mode A we can drop the condition of ultimate extinction and write instead*

$$\textit{birth-rate} = \gamma\,, \quad \textit{death-rate} = m\beta\,.$$

With the aid of this rather surprising result the discussion of the approximating

model becomes quite simple. For example, we can calculate the distribution of $\tilde{w} \equiv \tilde{Z}(\infty) - a$ and compare $E(\tilde{w})$ with Bailey's values for $E(w)$. If $m \leqq \rho$, then $E(\tilde{w})$ is the expected number of births in the appropriate birth-and-death process, so that

$$E(\tilde{w}) = \frac{ma}{\rho - m} \qquad (m \leqq \rho) . \tag{28}$$

We must therefore expect the approximating model to be quite unrealistic when $m = \rho$ and $E(\tilde{w}) = \infty$. If $m > \rho$, then

$$E(\tilde{w}) = \left(\frac{\rho}{m}\right)^a \frac{\rho a}{m - \rho} + \left[1 - \left(\frac{\rho}{m}\right)^a\right](\zeta - a) , \tag{29}$$

where ζ is the (unique) positive root of the equation

$$m + a - \zeta = me^{-\zeta/\rho} , \tag{30}$$

obtained by a straightforward application of (3). In table IV we compare $E(\tilde{w})$ with the values of $E(w)$ calculated by Bailey. No entries of $E(\tilde{w})$ are made when $m = \rho$ because then the approximation manifestly fails.

TABLE IV

COMPARISON OF $E(w)$ WITH $E(\tilde{w})$ (SHOWN IN PARENTHESES)

($a = 1$)

m	m/ρ					
	0.67	0.80	1	1.33	2	4
10	1.08 (2)	1.38 (4)	1.89	2.74 (3.8)	4.33 (4.8)	7.13 (7.5)
20	1.30 (2)	1.77 (4)	2.62	4.32 (5.1)	7.97 (8.8)	14.40 (14.9)
40	1.50 (2)	2.18 (4)	3.58	6.94 (7.4)	15.31 (16.8)	29.12 (29.6)

(N. T. J. Bailey's calculations.)

Once again the agreement in order of magnitude is quite as good as one could have hoped for apart from the breakdown when $m \cong \rho$ (that is, at the threshold itself).

5. The growth of a stochastic epidemic in time

The success of the approximating model in reproducing the main qualitative features of the original stochastic system when $m \neq \rho$ now suggests that we may be able to use it to gain some rough information about the growth of a stochastic epidemic in time. We should expect when $m > \rho$ that the sample epidemic curves will fall broadly into two classes: those (corresponding to mode A) which peter out fairly quickly, and those (corresponding to mode B) which approximate to the deterministic epidemic curve associated with the K and K equations (2). In order to test this conjecture we must appeal to Monte Carlo information. By a method described in detail elsewhere (Kendall, [7]) I have constructed a "random sample" of twenty artificial realisations of the McKendrick-Bartlett process ($X(t)$, $Y(t)$, $Z(t)$) with the following values for the parameters.

$$\beta = 1 , \quad \gamma = 10 , \quad \rho = 10 ,$$

TABLE V

LIFE HISTORIES OF 20 ARTIFICIAL REALIZATIONS OF A STOCHASTIC EPIDEMIC

($\beta = 1, \gamma = 10, \rho = 10; X(0) = 20, Y(0) = 1, Z(0) = 0$)

(In each main column the first entry gives the number of infectious persons in circulation at the stated epoch, and the second entry gives the number of cases "removed" during the stated interval.)

Epidemic	(a)		(b)		(c)		(d)		(e)		(f)		(g)		(h)		(i)		(j)		(k)		(l)		(m)		(n)		(o)		(p)		(q)		(r)		(s)		(t)	
Epoch	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ	Y	ΔZ
0.00	1		1		1		1		1		1		1		1		1		1		1		1		1		1		1		1		1		1		1	
		0		0		1		0		1		1		1		0		0		2		1		1		1		1		1		0		0		0		0
05	1		2		5		2		0		0		1		1		4		0		0		0		3		3		0		1		2		2		3		1	
		1		1		0		2						0		0		0								1		1				1		0		2		1	
10	0		1		9		0						3		1		7								3		3				0		2		2		7		1	
				0		6								1		1		2								3								4		1		2		2
15			1		5								3		3		11								3		3						7		1		7		3	
				0		1								3		0		5								0		1						2		0		3		0
20			1		5								1		4		9								4		2						9		1		11		6	
				1		2								2		0		3								1		1						5		1		2		3
25			1		3								0		4		7								9		2						6		0		10		6	
				1		2										4		3								4		2						1				2		2
30			5		3										4		5								6		1						5				9		9	
				4		1										1		2								4		0						3				3		4
35			6		2										9		3								2		1						3				6		6	
				2		1										3		1								1		1						0						4	
40			4		1										7		3								1		1						4				2		3	
				1		1										3		1								2		1						0				0		1
45			5		0										5		4								1		0						4				2		3	
				2												3		1								0								1				1		2
0.50			6												4		3								1								3				1		1	
				2												1		1								0								2				0		0
55			4												3		2								1								1				1		1	
				2												1		0								1								0				1		0
60			3												3		2								0								2				0		1	
				2												0		0																0						0
65			1												4		2																2						1	
				0												1		1																1						0
70			1												3		1																1						1	
				1												0		0																1						0
75			0												3		1																0						1	
																1		0																						0
80															2		1																						1	
																1		0																						0
85															1		1																						1	
																0		0																						0
90															1		1																						1	
																0		0																						0
95															1		1																						1	
																1		1																							1
1.00															0		0																						0	

and with the following initial values,

$$X(0) \equiv m = 20\,, \quad Y(0) \equiv a = 1\,, \quad Z(0) = 0\,.$$

It will be observed that $m/\rho = 2$, so that we may reasonably expect modes A and B to be well separated. The life-histories of the twenty epidemics are summarised in table V for the benefit of those who may wish to try out estimation techniques. For reasons of economy the data have been presented as concisely as possible. It will be noticed that $Z(t)$ can be calculated for each multiple of $\Delta t = 0.05$ by cumulating the entries of ΔZ, and that $X(t)$ is then determined by the conservation law $X + Y + Z = 21$. The entries in the columns headed ΔZ, when plotted against the time, constitute the empirical notifications "curve," while the graph of Y against t is the epidemic curve. There is of course no reason why Y and ΔZ should remain in a constant proportion for a *stochastic* epidemic.

We can check the representativeness of this random sample of 20 stochastic epidemics by examining the empirical distribution of the number of secondary cases and then comparing it with the theoretical distribution determined by Bailey. The result of the comparison is shown in table VI.

TABLE VI

COMPARISON OF EMPIRICAL AND THEORETICAL DISTRIBUTIONS OF w

Values of w	0-2	3-17	18-20
Observed (Kendall)	9	5	6
Expected (Bailey)	9.0	7.2	3.8

The method of grouping has been chosen so as to isolate the two peaks at the extremities of the range. The value of χ^2 (with 2 degrees of freedom) is 1.95, which is satisfactory so far as it goes.

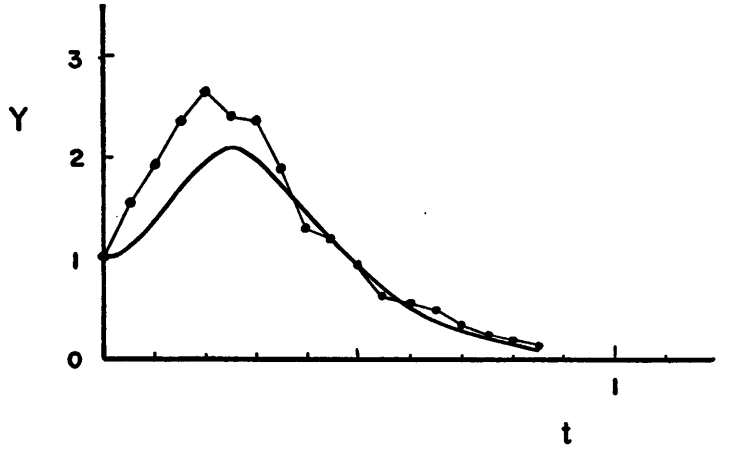

FIGURE 1

The epidemic curve (the number of infectious individuals in circulation, plotted against the time). The arithmetic mean of a sample of 20 artificial realizations of $Y(t)$ (linked spots) is compared with $E\{\tilde{Y}(t)\}$ (continuous curve).

We next invite the reader to examine table VII, which compares (x, y, z) for the associated K and K deterministic epidemic (exact solution as given in section 3) with the arithmetic means of (X, Y, Z) for the twenty artificial epidemics, for various epochs t. It will be seen that the two epidemic curves have notably different peak heights and this is not surprising; Feller [5] has remarked that in such non-linear systems the stochastic mean will not normally coincide with the deterministic equivalent.

TABLE VII

A Deterministic Epidemic and the (Empirical) Expectation Behaviour of its Stochastic Counterpart

(For details, see text. The values of x, y and z have been obtained from the exact formulas of section 3; the values of $\bar{X}$, $\bar{Y}$ and $\bar{Z}$ are the arithmetic means of X, Y, and Z for 20 artificial stochastic epidemics.)

Epoch	x	$\bar{X}$	y	$\bar{Y}$	z	$\bar{Z}$
0.00	20.0	20.0	1.0	1.0	0.0	0.0
05	18.8	18.9	1.6	1.6	0.6	0.6
10	16.9	17.9	2.4	2.0	1.7	1.2
15	14.8	16.2	3.2	2.4	3.0	2.5
20	12.4	15.2	3.8	2.7	4.8	3.2
25	10.2	14.4	4.1	2.4	6.7	4.3
30	8.4	13.4	3.9	2.4	8.7	5.3
35	7.0	12.7	3.4	1.9	10.6	6.4
40	5.9	12.5	2.9	1.3	12.2	7.3
45	5.2	12.1	2.3	1.2	13.5	7.8
50	4.7	11.8	1.8	1.0	14.5	8.3
55	4.3	11.8	1.4	0.7	15.3	8.6
60	4.1	11.7	1.0	0.6	15.9	8.8
65	3.9	11.6	0.8	0.5	16.3	8.9
70	3.7	11.6	0.6	0.4	16.7	9.1
75	3.7	11.6	0.4	0.3	16.9	9.2
80	3.6	11.6	0.3	0.2	17.1	9.2
85	3.6	11.6	0.2	0.2	17.2	9.3
∞	3.5	11.6	0.0	0.0	17.5	9.4

It will now be appropriate to compare the empirical mean stochastic epidemic curve of table VII (the graph of $\bar{Y}(t)$ against t) with the calculated mean of $\tilde{Y}(t)$ (the corresponding random variable in the approximating system). The comparison is shown in figure 1; the curve gives the calculated mean value of $\tilde{Y}(t)$ and the linked spots locate empirical values of $\bar{Y}(t)$.

The agreement so far as the post-peak behaviour is concerned is most satisfactory, and if it is confirmed by a more extensive series of Monte Carlo experiments then we can claim to have constructed a rough approximation to the expectation behaviour of a stochastic epidemic.

It must be realised, however, that the *mean* stochastic epidemic curve when taken by itself gives no indication of the large fluctuations which may be shown by individual epidemics. In order to illustrate this we show in figure 2 (α, β, γ) the 20 individual epidemic curves, plotted as if observed at regularly spaced epochs, and

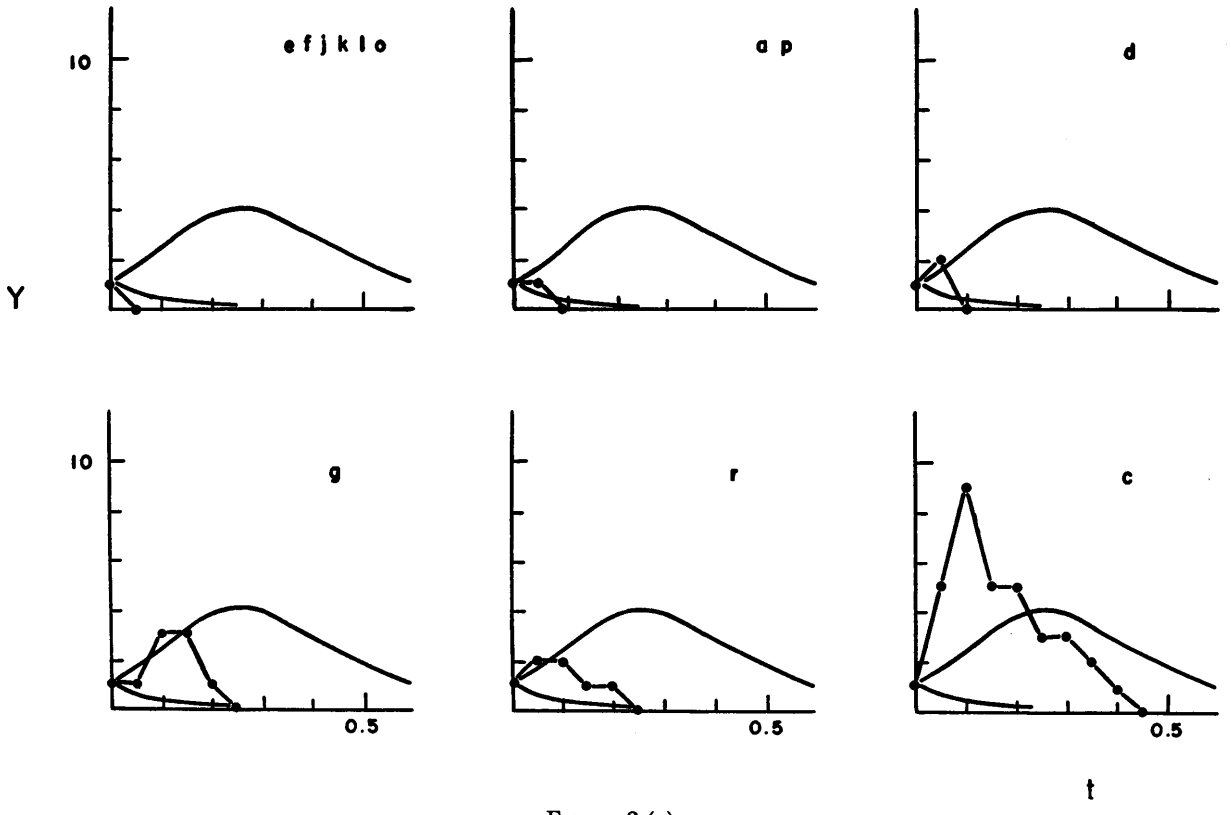

FIGURE 2 (*a*)

The epidemic curves for twelve artificial stochastic epidemics (linked spots). The continuous curves show the conditional means for mode A (J-shaped) and mode B (peaked).

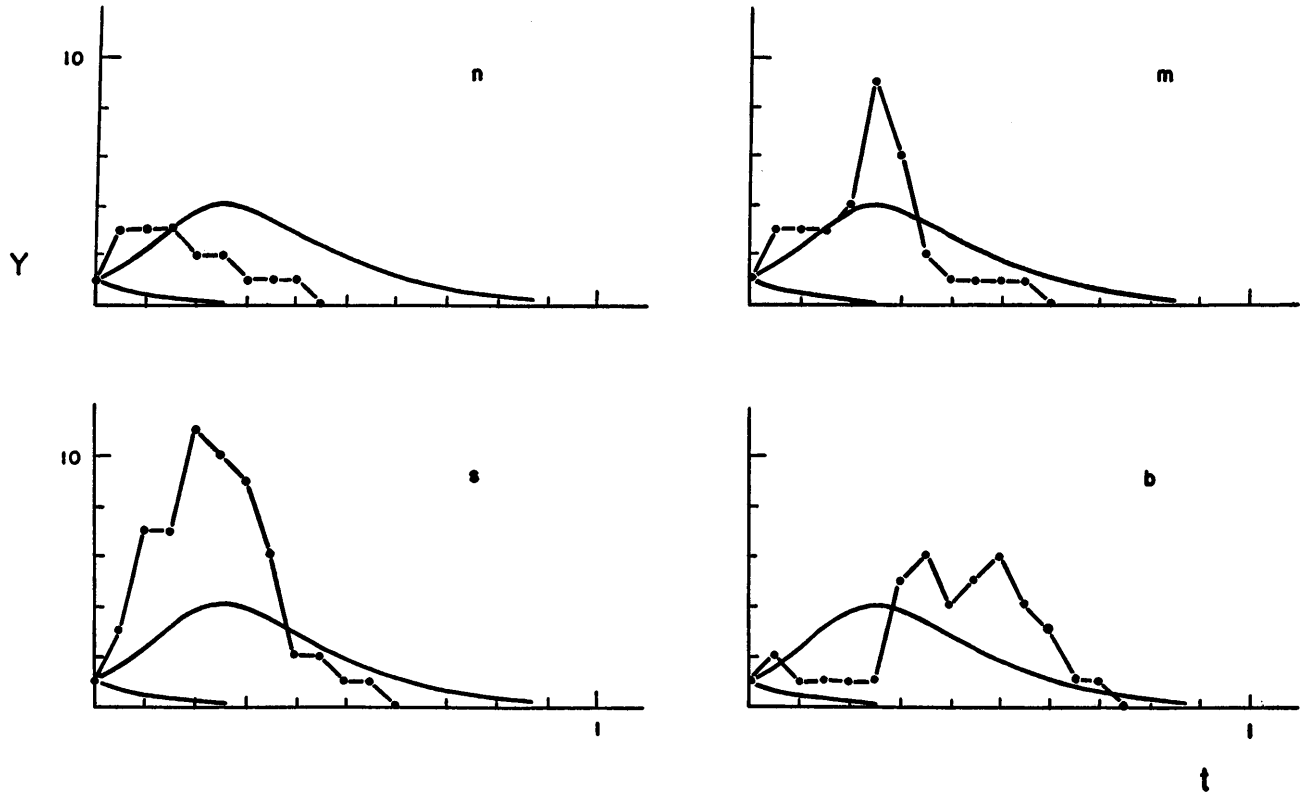

FIGURE 2 (β)
The epidemic curves for four artificial stochastic epidemics.

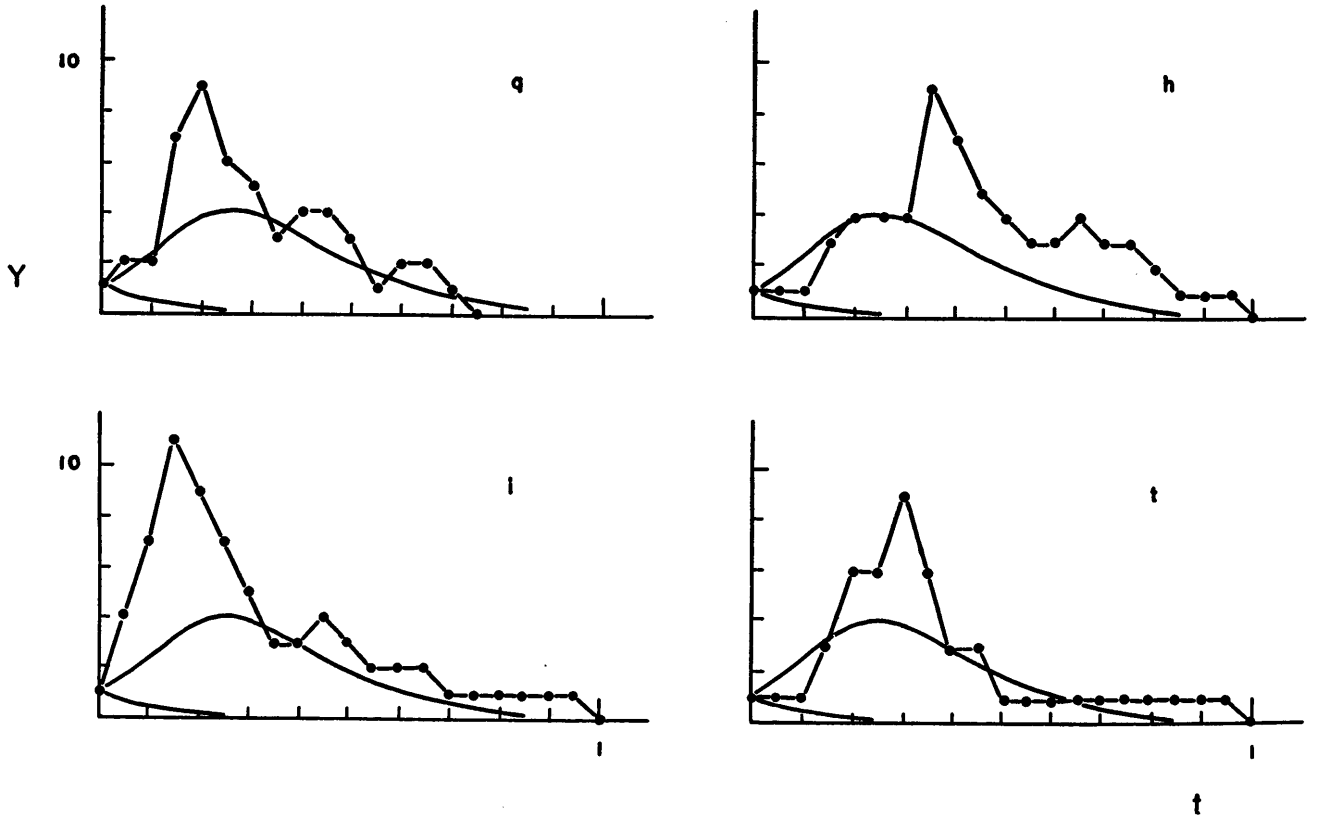

FIGURE 2 (γ)
The epidemic curves for four artificial stochastic epidemics.

on each figure we have drawn for comparison the conditional mean epidemic curves for the behaviour of the approximating system in each of the modes A and B. (The peaked curve gives the conditional mean for mode B while the J-shaped curve relates to mode A). The violent fluctuations are very striking. Epidemics *e*, *f*, *j*, *k*, *l*, *o*; *a*, *p*; *d*; *g*; and *r* (on figure 2(α)) may be said to display mode A behaviour while epidemics *c*, *n*, *m*, *s*, *b*, *q*, *h*, *i* and *t* display mode B behaviour. The numbers in the two classes are 11 and 9, respectively. It will be noted that with the given values of the constants, pr (A) = pr (B) = $\frac{1}{2}$, so that equal frequencies for the two modes would be expected.

In conclusion I should like to thank Mr. N. T. J. Bailey, Professor M. S. Bartlett, Mr. J. Hajnal, Mr. G. E. H. Reuter and Dr. W. A. O'N. Waugh for helpful discussions in connexion with the present work. Professor Bartlett has developed independently and rather differently the idea of using an approximating stochastic system; details of his work will be found in his book [4] and in his contribution to this Symposium.

REFERENCES

[1] N. T. J. Bailey, "The total size of a general stochastic epidemic," *Biometrika*, Vol. 40 (1953), pp. 177–185.

[2] ———, "Some problems in the statistical analysis of epidemic data," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 17 (1955), pp. 35–68.

[3] M. S. Bartlett, "Some evolutionary stochastic processes," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 11 (1949), pp. 211–229.

[4] ———, *An Introduction to Stochastic Processes*, Cambridge, The University Press, 1955.

[5] W. Feller, "Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung," *Acta Biotheoretica*, Vol. 5 (1939), pp. 11–40.

[6] F. G. Foster, "A note on Bailey's and Whittle's treatment of a general stochastic epidemic," *Biometrika*, Vol. 42 (1955), pp. 123–125.

[7] D. G. Kendall, "An artificial realization of a simple birth-and-death process," *Jour. Roy. Stat. Soc.*, Ser. B, Vol. 12 (1950), pp. 116–119.

[8] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proc. Roy. Soc. Lond.*, Ser. A, Vol. 115 (1927), pp. 700–721.

[9] A. G. McKendrick, "Applications of mathematics to medical problems," *Proc. Edinburgh Math. Soc.*, Vol. 44 (1926), pp. 1–34.

[10] W. A. O'N. Waugh, "Conditioned Markov processes," (in preparation).

[11] P. Whittle, "The outcome of a stochastic epidemic—a note on Bailey's paper," *Biometrika*, Vol. 42 (1955), pp. 116–122.