Biometrika (1980), **67**, 1, *pp.* 191–8
Printed in Great Britain

On the spread of a disease with gamma distributed latent and infectious periods

BY DOROTHY ANDERSON
Department of Statistics, Australian National University, Canberra

AND RAY WATSON
Department of Statistics, University of Melbourne

SUMMARY

In this paper we investigate a model with assumptions as for the so-called general epidemic model, except that the latent and infectious periods are independently gamma distributed. For this model, we obtain a number of results concerning the progress of the epidemic. These results allow some assessment of the effects of the latent period and the infectious period on the behaviour of the epidemic model. Further we note that many of the results can be generalized.

Some key words: Infectious period; Latent period; Size distribution; Stochastic model; Threshold.

1. INTRODUCTION

The so-called general epidemic model, which we will call the standard epidemic model, assumes a zero latent period and an exponentially distributed infectious period. Little work has been done on the effect on the behaviour of the model of modification of these unrealistic assumptions. Bailey (1964) proposed a model of the type considered here, but derived only a deterministic large-population approximation.

We investigate a model, denoted by $\mathscr{P}_{lm}$ ($l = 0, 1, \ldots$, $m = 1, 2, \ldots$), with assumptions as for the standard epidemic model except that the latent periods and infectious periods have independent gamma distributions with densities of the form

$$f_{k\phi}(t) = e^{-k\phi t} t^{k-1} (k\phi)^k \{\Gamma(k)\}^{-1} \quad (t > 0), \tag{1}$$

for $k = 1, 2, \ldots$ and $\phi > 0$, which has mean ϕ^{-1} and a variance $k^{-1}\phi^{-2}$. If $k = 0$ we define the distribution to be degenerate at zero. We note that the family of distributions $\{f_{k\phi}\}$ includes as special cases the exponential distribution, $k = 1$, distributions which are close to normal, k large, and the degenerate distribution at ϕ^{-1}, as $k \to \infty$. The latent period has density $f_{l\alpha}$ and the infectious period density $f_{m\gamma}$. The model $\mathscr{P}_{01}$ is in fact the standard epidemic model.

For the model $\mathscr{P}_{lm}$, we obtain results concerning the behaviour of the path of the process by means of the deterministic approximation and a martingale central limit theorem result. Equations for the distribution of the size of the outbreak are specified and approximations are derived for the probability of a minor outbreak, the distribution of the size of a minor outbreak and the distribution of the size of a major outbreak. Further, we obtain an approximation for the initial rate of increase of the number of infected individuals and hence the mean time for the epidemic to reach its peak.

These results give a reasonably detailed description of the behaviour of the model $\mathscr{P}_{lm}$, allowing an assessment of the effects of the parameters l and m, and also enabling a comparison of $\mathscr{P}_{lm}$ with the standard epidemic model $\mathscr{P}_{01}$ to be made.

Finally many of the results derived for the model $\mathscr{P}_{lm}$ can be extended to a more general model.

2. Definition of the model

Let $X(t)$ denote the number of susceptible individuals at time t. Let $Y_i'(t)$ denote the number of latent infected individuals in the ith stage at time t ($i = 1, \ldots, l$), with $Y'(t) = \Sigma\, Y_i'(t)$, the number of latent infected individuals at time t. We note that if $l = 0$ then $Y'(t) = 0$. Let $Y_j''(t)$ denote the number of infective individuals in the jth stage at time t ($j = 1, \ldots, m$), with $Y''(t) = \Sigma\, Y_j''(t)$ the number of infective individuals at time t. Then $Y(t) = Y'(t) + Y''(t)$ is the number of infected individuals at time t. Let $Z(t)$ denote the number of removed individuals at time t.

The process $\mathscr{P}_{lm} = \{P(t), t \geqslant 0\}$, where $P(t) = \{X(t), Y_1'(t), \ldots, Y_l'(t), Y_1''(t), \ldots, Y_m''(t), Z(t)\}$, is a continuous time Markov process with initial condition

$$P(0) = (n, a_1', \ldots, a_l', a_1'', \ldots, a_m'', 0) \tag{2}$$

and we let $a' = \Sigma a_i'$, $a'' = \Sigma a_j''$ and $a = a' + a''$. The transition probabilities for the process $\mathscr{P}_{lm}$ are given by

$$\operatorname{pr}\{P(t+dt) = p - e(x) + e(y_1') \mid P(t) = p\} = n^{-1}\beta x \sum_{j=1}^{m} y_j''\, dt, \tag{3a}$$

$$\operatorname{pr}\{P(t+dt) = p - e(y_i') + e(y_{i+1}') \mid P(t) = p\} = l\alpha y_i'\, dt \quad (i = 1, \ldots, l-1), \tag{3b}$$

$$\operatorname{pr}\{P(t+dt) = p - e(y_l') + e(y_1'') \mid P(t) = p\} = l\alpha y_l'\, dt, \tag{3c}$$

$$\operatorname{pr}\{P(t+dt) = p - e(y_j'') + e(y_{j+1}'') \mid P(t) = p\} = m\gamma y_j''\, dt \quad (j = 1, \ldots, m-1), \tag{3d}$$

$$\operatorname{pr}\{P(t+dt) = p - e(y_m'') + e(z) \mid P(t) = p\} = m\gamma y_m''\, dt, \tag{3e}$$

where $p = (x, y_1', \ldots, y_l', y_1'', \ldots, y_m'', z)$, and $e(p_k)$ denotes a vector of $l+m+1$ zeros and a single one in the position corresponding to the component p_k of the vector p. In the case $l = 0$, the transition probabilities (3a), (3b) and (3c) are modified in the obvious way.

The mechanism of the process $\mathscr{P}_{lm}$ is thus quite simple. If $l > 0$, each infected individual passes through l stages before becoming infective, spending an exponentially distributed time in each stage which is independent of time spent in every other stage. Thus, the latent period is the sum of l independent exponentially distributed random variables each having mean $(l\alpha)^{-1}$. If $l = 0$, an infected individual immediately becomes infective so that the latent period is zero. On becoming infective, each individual passes through m stages such that at each stage he is equally capable of spreading the infection. The infectious period is thus the sum of m independent exponential random variables each having mean $(m\gamma)^{-1}$.

We define the parameter $\theta = \beta/\gamma$ which is most important in describing the behaviour of the process $\mathscr{P}_{lm}$, as it is for the standard epidemic model.

We observe that transitions cease when $Y(t) = 0$, and we let $T = \inf\{t\colon Y(t) = 0\}$, so that T denotes the duration of the epidemic. The final state of the process is denoted by

$$P(T) = (n - S, 0, \ldots, 0, 0, \ldots, 0, S + a),$$

where S denotes the size of the outbreak.

3. Deterministic approximation

The deterministic approximation for the process $\mathscr{P}_{lm}$, namely

$$p(t) = \{x(t), y_1'(t), \ldots, y_l'(t), y_1''(t), \ldots, y_m''(t), z(t)\}$$

is such that

$$\frac{dx}{dt} = n^{-1}\beta x y'', \quad \frac{dy_1'}{dt} = n^{-1}\beta x y'' - l\alpha y_1', \quad \frac{dy_i'}{dt} = l\alpha y_{i-1}' - l\alpha y_i' \quad (i = 2, \ldots, l),$$

$$\frac{dy_1''}{dt} = l\alpha y_l' - m\gamma y_1'', \quad \frac{dy_j''}{dt} = m\gamma y_{j-1}'' - m\gamma y_j'' \quad (j = 2, \ldots, m), \quad \frac{dz}{dt} = m\gamma y_m'',$$

where $y''(t) = \Sigma y_j''(t)$, and with initial conditions as specified by (2). We define

$$y^*(t) = y'(t) + m^{-1}\sum_{j=1}^{m}(m-j+1)\,y_j''(t), \quad z^*(t) = m^{-1}\sum_{j=1}^{m}(j-1)\,y''(t) + z(t),$$

where $y'(t) = \Sigma y_i'(t)$, and

$$y^*(0) = a^* = a' + m^{-1}\sum_{j=1}^{m}(m-j+1)\,a_j'', \quad z^*(0) = b^* = m^{-1}\sum_{j=1}^{m}(j-1)\,a_j'',$$

so that $a^* + b^* = a$. It is readily shown that

$$k(t) = -n\log\{n/x(t)\} + \theta\{z^*(t) - b^*\} \tag{4}$$

is a constant of the motion, and using the initial conditions (2), we deduce that $k(t) = 0$ for all $t \geqslant 0$. Thus, letting $t \to \infty$, we obtain the following equation for the deterministic approximation to the severity of the outbreak, $\sigma = 1 - n^{-1}x(\infty)$:

$$\log(1-\sigma) + \theta(\sigma + \alpha^*) = 0, \tag{5}$$

where $\alpha^* = n^{-1}a^*$, which is of the same form as the severity equation for the standard epidemic model with $y(0) = a^*$. We note that if all the initial infectives are fresh, i.e. in one of the latent stages of infection or in the first infective stage, then $a^* = a$, but, otherwise, $a^* < a$. In any application $n^{-1}a$ will be small so that the difference between a^* and a will have little effect.

Also, it follows from (4) that

$$y^*(t) = n + a^* - x(t) - \theta^{-1}n\log\{n/x(t)\}.$$

Thus, the relationship between y^* and x is identical to that between y and x for the standard epidemic model with $y(0) = a^*$. We let $\mathscr{P}_{01}^*$ denote the standard epidemic model with $Y(0)$ equal to the smallest integer not less than a^*. This is the most reasonable choice for a corresponding epidemic model to allow for the fact that for the model $\mathscr{P}_{lm}$ some of the initial infectives are less than fully effective in spreading the disease.

Now, since $y^* < y$, it follows that the number of infected individuals in the population is, at all corresponding stages of the outbreak, greater for $\mathscr{P}_{lm}$ than it is for $\mathscr{P}_{01}^*$; i.e. as indicated in Fig. 1, the deterministic path of the process $\mathscr{P}_{01}^*$ lies entirely under that of $\mathscr{P}_{lm}$. However, we note that, as indicated in § 9, these paths need not be traversed at the same rate, and if $l > 0$ there may be a substantial difference in the rates.

4. Equations for the size distribution

If we let $\pi_i = \mathrm{pr}(S = i)$ $(i = 0, 1, \ldots, n)$ then, following the method of Watson (1980c), we can easily show that

$$\sum_{i=0}^{k}\pi_i\frac{k^{(i)}}{n^{(i)}}\{1 + m^{-1}\theta(1 - n^{-1}k)\}^{m(a+i)} = 1 \quad (k = 0, 1, \ldots, n).$$

Thus the size distribution depends on m but is independent of l. That this should be so can

also be seen from the fact that, independently of l, each susceptible has potentially exactly the same amount of exposure to infection; only the time scale is altered.

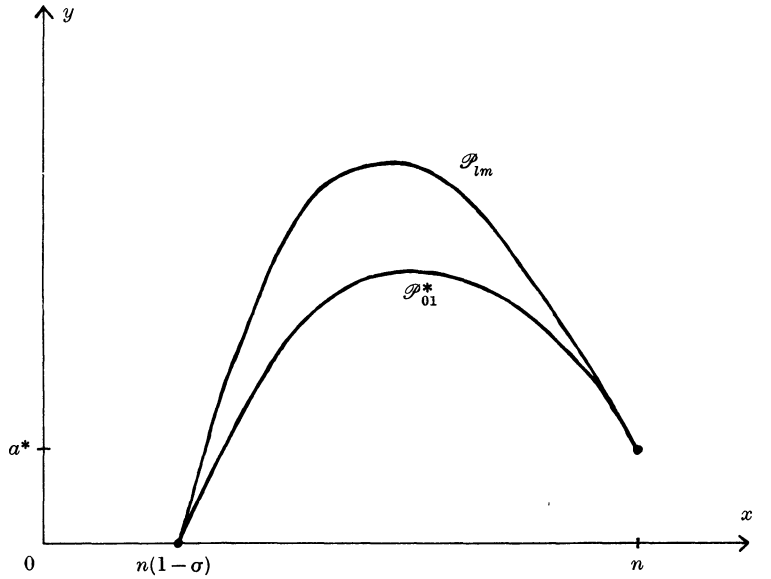

Fig. 1. Deterministic approximations for $\mathscr{P}_{lm}$ and for the corresponding standard epidemic model $\mathscr{P}^*_{01}$.

5. BRANCHING PROCESS APPROXIMATION FOR MINOR OUTBREAKS

We can approximate the initial behaviour of the number of infectives by assuming that the number of susceptibles remains fixed at its initial value. This approximation is valid for the entire course of a minor outbreak, and so leads to an approximation for the probability of a minor outbreak and also for the distribution of the size of a minor outbreak.

Thus, we can approximate the development of the process $\{Y(t), t \geqslant 0\}$ by a branching process with offspring probability generating function $Q(\xi) = \{1 + m^{-1}\theta(1-\xi)\}^{-m}$. However, it is more useful to consider the process $\{mY^*(t), t \geqslant 0\}$, where

$$mY^*(t) = mY'(t) + \Sigma\,(m-j+1)\,Y''_j(t).$$

We note that $mY^*(t)$ can be interpreted as the number of infective stages remaining to the infected individuals in the population at time t. Thus, the development of the process $\{mY^*(t), t \geqslant 0\}$ can be approximated by a branching process with offspring probability generating function

$$Q^*(\xi) = \{1 + m^{-1}\theta(1-\xi^m)\}^{-1},$$

which has mean θ. Thus, this approximation indicates that a major outbreak is possible only if $\theta > 1$, in which case it occurs with probability $1 - \pi^{a^*}$, where π is the smaller root of

$$\pi\{1 + m^{-1}\theta(1-\pi)\}^m = 1. \tag{6}$$

From (6), we see that π is independent of l, and that, for given θ, π is a decreasing function of m, that is $\pi = \pi(m, \theta)$ is such that

$$\pi(\infty, \theta) = 1 - \varepsilon(\theta, 0) \leqslant \pi(m, \theta) \leqslant \pi(1, \theta) = \theta^{-1},$$

where $\varepsilon(\theta, \alpha^*)$ denotes the positive solution of (5). The behaviour of $\pi(m, \theta)$ is indicated in Fig. 2.

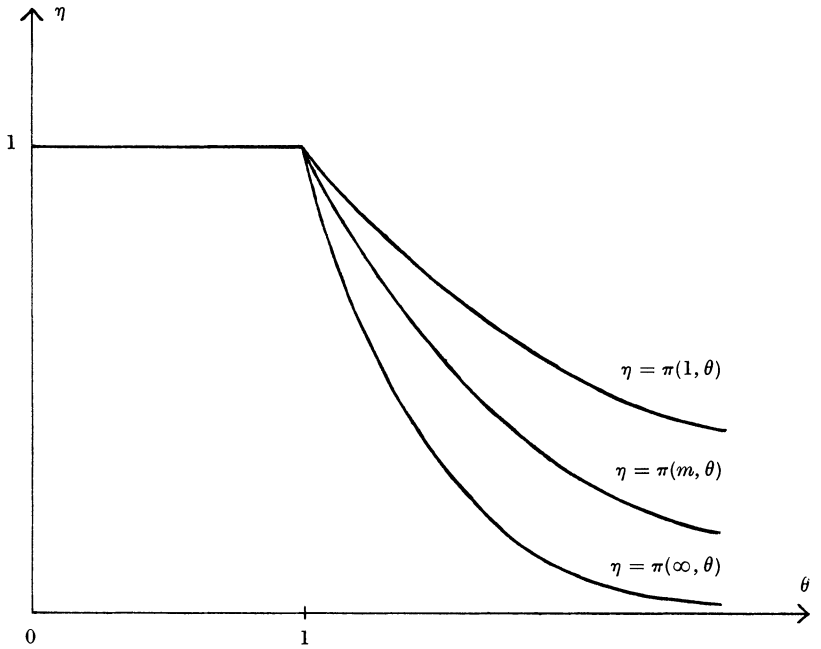

Fig. 2. Behaviour of $\pi(m, \theta)$, the approximation for the probability of a minor outbreak initiated by one fresh infective.

If S_0 denotes the size of a minor outbreak, then it follows from the above, if we use a result of Waugh (1958), that the probability generating function of S_0, $P_{S_0}(\xi) \doteq P(\xi)^{a^*}$, where P is such that

$$P(\xi) = [1 + m^{-1}\phi\{1 - \xi P(\xi)\}]^{-m}, \tag{7}$$

where $\phi = \theta\pi^{1+1/m}$ and π is given by (6). Further, it is then easily deduced that approximations for the mean and variance of a minor outbreak are given by

$$E(S_0) \doteq a^*\phi(1-\phi)^{-1}, \quad \text{var}\,(S_0) \doteq a^*\phi(1 + m^{-1}\phi)\,(1-\phi)^{-3}.$$

The approximations derived from the branching process are less worthwhile for values of θ near 1, i.e. for a specified accuracy, larger values of n are required for θ nearer 1.

6. Martingale central limit theorem approximation for major outbreak

If we define

$$Z^*(t) = m^{-1}\sum_{j=1}^{m}(j-1)\,Y_j''(t) + Z(t), \quad \Delta Z^*(t) = Z^*(t) - b^*,$$

and let

$$K(t) = -n\nabla H_1\{X(t)\} + \theta\Delta Z^*(t),$$

where

$$H_\nu(x) = \sum_{i=1}^{x} i^{-\nu} \quad (x = 1, 2, \ldots), \quad H_\nu(0) = 0, \quad \nabla H_\nu(x) = H_\nu(n) - H_\nu(x),$$

then $\{K(t), t \geqslant 0\}$ is a zero-mean square-integrable martingale with respect to $\{F(t), t \geqslant 0\}$, where $F(t)$ denotes the σ-field generated by $\{P(\tau), 0 \leqslant \tau \leqslant t\}$.

For the standard epidemic model, Watson (1980b) has shown that a central limit theorem holds for the martingale $\{K(t), t \geqslant 0\}$, such that, in the event of a major outbreak, as $n \to \infty$ we have in distribution that

$$\frac{-n\nabla H_1\{X(t)\} + \theta \Delta Z(t)}{\sqrt{[n^2 \nabla H_2\{X(t)\} + \theta^2 \Delta Z(t)]}} \to \mathcal{N},$$

where $\mathcal{N}$ denotes the standard normal distribution. The proof carries over to the present model on replacing $Z(t)$ by $Z^*(t)$. Thus, we find that, for large values of n, in the event of a major outbreak

$$\frac{-n\nabla H_1\{X(t)\} + \theta \Delta Z^*(t)}{\sqrt{[n^2 \nabla H_2\{X(t)\} + m^{-1}\theta^2 \Delta Z^*(t)]}} \doteq \mathcal{N}.$$

If S_1 denotes the size of a major outbreak, i.e. the size conditional on the occurrence of a major outbreak, then, since at the end of the outbreak $X = n - S_1$ and $\Delta Z^* = S_1 + a^*$, we deduce that

$$\frac{-n\nabla H_1(n - S_1) + \theta(S_1 + a^*)}{\sqrt{\{n^2 \nabla H_2(n - S_1) + m^{-1}\theta^2 (S_1 + a^*)\}}} \doteq \mathcal{N}. \tag{8}$$

It is simple to derive from (8) an approximation for the distribution of the size of a major outbreak. Further, from (8), we find that

$$E(S_1) \doteq n\sigma + 2gv_m(1-\sigma)^{-2}, \quad \text{var}\,(S_1) \doteq ng^2 v_m,$$

where $g = (1-\sigma)^{-1} - \theta > 0$ and $v_m = \sigma(1-\sigma)^{-1} + m^{-1}\theta^2(\sigma + n^{-1}a^*)$. Since v_m is a decreasing function of m, the above approximations suggest that both the mean and variance of the size of a major outbreak decrease as m increases.

7. Approximation for the size distribution

The size distribution can now be approximated using the above results, since the size distribution can be regarded as a mixture of the minor outbreak distribution approximated by (7) and the major outbreak distribution approximated by (8) with mixing probability approximated by (6).

Figure 3 compares the approximation and the empirical distribution obtained from the simulation of 10,000 outbreaks for $l = 0$, $m = 1, 2, 5, \infty$ with $n = 100$, $a_1'' = 1$, $a_j'' = 0$ $(j > 1)$ and $\theta = 2$.

8. Inference based on observation of the size of the outbreak

We note that these approximate results for the size distribution enable inference on θ based on observation of the size when m is assumed known. This would occur when m, or even m and γ, are specified by the known behaviour of the infectious period. In the event of a minor outbreak, standard branching process results for inference on the mean of the offspring distribution can be used for inference on θ.

In the event of a major outbreak, the result (8) suggests an estimator

$$\tilde{\theta} = n\nabla H_1(n - S_1)/(S_1 + a^*),$$

where S_1 denotes the size of the major outbreak, for which

$$\frac{(\bar{\theta}-\theta)(S_1+a^*)}{\sqrt{\{n^2\nabla H_2(n-S_1)+m^{-1}\theta^2(S_1+a^*)\}}} \doteq \mathcal{N}.$$

For the standard epidemic model, $\bar{\theta}$ is the maximum likelihood estimator of θ based on observation of the size only, and $\bar{\theta}$ is remarkably efficient compared to the maximum likelihood estimator based on total observation of the process (Watson, 1980a). It is easily shown that these characteristics carry over to the present model.

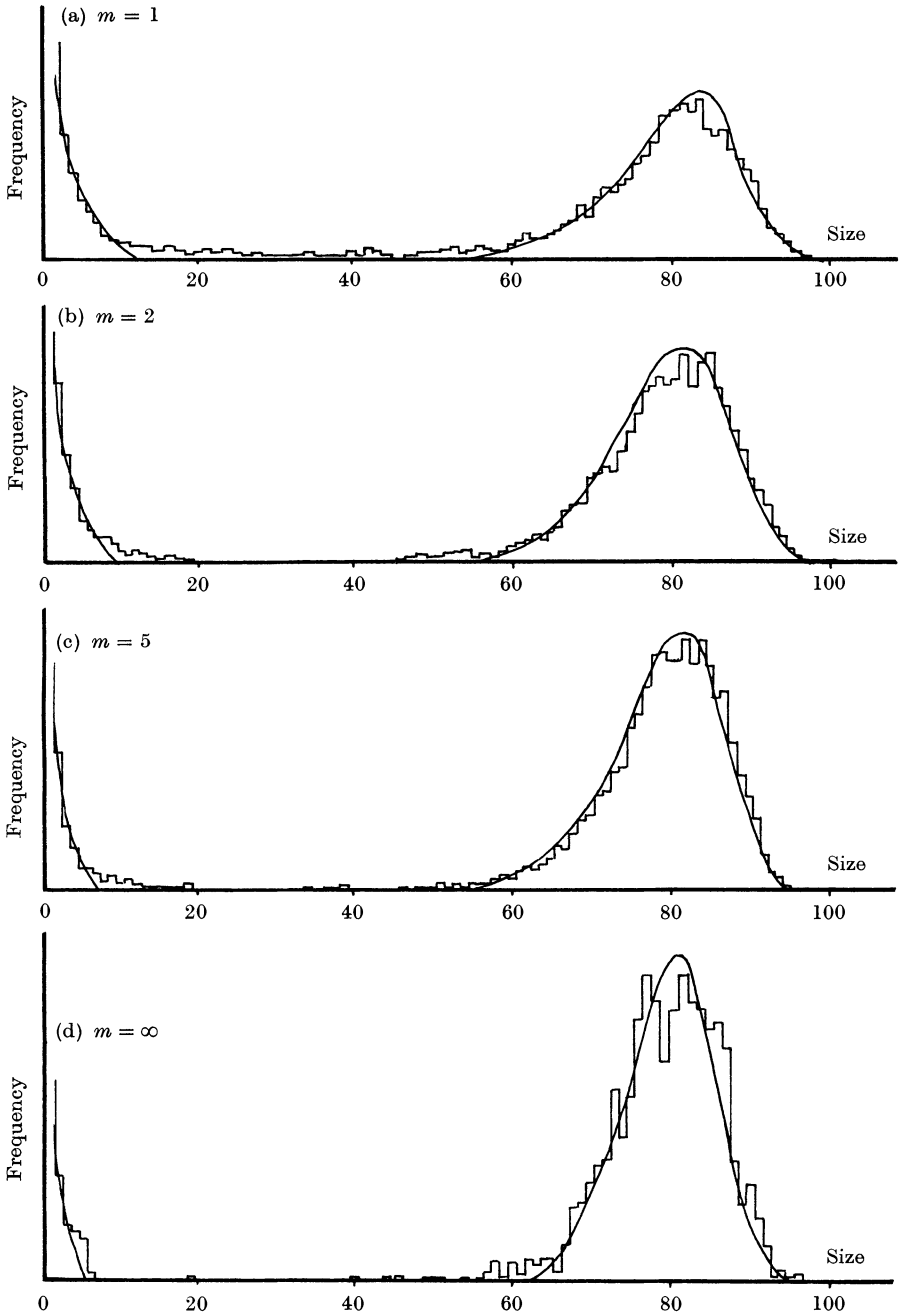

Fig. 3. Empirical size distribution for $n = 100$, $\theta = 2$ and $m = 1, 2, 5$ and ∞ based on simulation of 10,000 outbreaks, compared with the approximation of § 7.

9. APPROXIMATION FOR THE MEAN CRITICAL TIME

The critical time of the epidemic is defined as the time for a major outbreak to reach its peak. It is easily seen that the critical time is asymptotically equal to the time taken for the process to move away from the neighbourhood of the set of states for which the number of infectives is zero. Thus, asymptotically, it is determined by the initial rate of increase of the number of infected individuals. Again, therefore, we need to consider the behaviour of the process $\mathscr{P}_{lm}$ in the initial stages of the outbreak.

If we again assume that the number of susceptibles remains fixed at its initial value, then the vector process $\{P_Y(t), t \geqslant 0\}$, where $P_Y(t) = \{Y_1'(t), \ldots, Y_l'(t), Y_1''(t), \ldots, Y_m''(t)\}$, is approximated by a multivariate birth and death process. This approximation is valid in the initial stages of an outbreak and improves as n increases. Thus we find that $E\{P_Y(t)\} \doteq \mu(t)$, where $\mu(t) = \{\mu_1'(t), \ldots, \mu_l'(t), \mu_1''(t), \ldots, \mu_m''(t)\}$ is such that $d\mu/dt = \Omega\mu$, where the matrix Ω has a simple form with elements $\pm l\alpha$, $\pm m\gamma$, β and 0. Let ω denote the largest eigenvalue of Ω; then ω is the largest root of

$$\left(1+\frac{\lambda}{m\gamma}\right)^m\left\{1-\frac{\lambda}{\beta}\left(1+\frac{\lambda}{l\alpha}\right)^l\right\} = 1. \tag{9}$$

Now $\omega = \omega_{lm}$ is a measure of the initial rate of increase of the process $\{Y(t), t \geqslant 0\}$, and the mean critical time is such that

$$\tau_{lm} \doteq \omega_{lm}^{-1} \log n.$$

When $l = 0$, comparison of (9) with (6) indicates that if $\theta > 1$, then

$$\omega_{0m} = \beta\{1 - \pi(m, \theta)\},$$

where $\pi(m, \theta)$ is the probability of a minor outbreak.

10. GENERALIZATIONS

Many of the above results can be generalized (i) by allowing variable rates of progression through the stages, and (ii) by allowing the number of latent stages and the number of infective stages to be random. The simplest case of such randomization is the carrier-borne epidemic model of Downton (1968) for which $\text{pr}(m = 0) = 1 - \pi$, $\text{pr}(m = 1) = \pi$.

Specifying and working with a general model of this form is very cumbersome, but the deterministic analysis and the branching process approximation hold with minor modifications, though we can no longer allow for initial nonfresh infectives. The martingale result can be extended and the matrix corresponding to Ω is easily specified from which the initial rate of increase of the number of infective individuals can be assessed.

REFERENCES

BAILEY, N. T. J. (1964). Some stochastic models for small epidemics in large populations. *Appl. Statist.* **13**, 9–19.

DOWNTON, F. (1968). The ultimate size of carrier-borne epidemics. *Biometrika* **55**, 277–89.

WATSON, R. K. (1980a). A useful random time-scale transformation for the standard epidemic model. *J. Appl. Prob.* To appear.

WATSON, R. K. (1980b). An application of a martingale central limit theorem to the standard epidemic model. *Stoc. Proc. Appl.* To appear.

WATSON, R. K. (1980c). On the size distribution for some epidemic models. *J. Appl. Prob.* To appear.

WAUGH, W. A. O'N. (1958). Conditioned Markov processes. *Biometrika* **45**, 241–9.

[*Received July* 1979. *Revised October* 1979]