Epidemics and Rumours: A Survey

Author(s): Klaus Dietz

Source: *Journal of the Royal Statistical Society. Series A (General)*, 1967, Vol. 130, No. 4 (1967), pp. 505-528

Published by: Wiley for the Royal Statistical Society

Stable URL: https://www.jstor.org/stable/2982521

Epidemics and Rumours: A Survey

By KLAUS DIETZ

Institut für medizinische Statistik und Dokumentation, Universität Freiburg i. Br.

SUMMARY

In this paper we review the recent mathematical contributions to the description of the spread of epidemics and rumours. The most important concepts of epidemiology are defined and the various models are classified according to their assumptions.

1. INTRODUCTION

SINCE 1957, the date of publication of Bailey's book on the mathematical theory of epidemics, the contributions to this subject have themselves behaved like an epidemic; a review of recent work in this field now seems to be desirable. In spite of lively progress on the part of theory, the application of models to real epidemic situations is still at its beginning.

We are concerned with deterministic and stochastic models for the spread of some "infectious" or "contagious" phenomenon through a population as a function of time. This comprises not only the transmission of a pathogenic agent from an infected host to an uninfected susceptible, with or without an intermediate animal vector, but also the propagation of ideas, rumours and consumers' goods. We do not deal with statistical methods for the detection of whether person-to-person infection is an important element in any given disease, e.g. in the case of leukaemia (see Barton and David, 1966; Barton *et al.*, 1965; Ederer *et al.*, 1964; Fix, 1967; Gaffey, 1954; Knox, 1964, 1965; Pyke, 1965; Yang, 1966). We also exclude from consideration the so-called catalytic models in epidemiology due to Muench (1959). He assumes a constant force of infection which acts on the population much as a catalyst acts on the aggregate. By means of simple differential equations he derives the age distribution of those who have recovered from some infectious disease which confers perpetual immunity, e.g. measles or yellow-fever. (The first mathematical approach to this problem in the case of smallpox appears to be the paper by Daniel Bernoulli, 1760!)

In Section 2 we define the most important concepts of epidemiology and classify the various epidemic models according to their assumptions. A description of recent contributions to these models is contained in Sections 3–8. Section 9 deals with some models for the spread of rumours. The references are to be considered as a supplement to those of Bailey (1957); for the literature up to 1957 the reader is referred to Bailey's book.

Previous reviews have been written by Serfling (1952), Sakino (1962a) about deterministic models, Maia (1952) about chain-binomial models, Taylor (1956), Bartlett (1955, Chap. 4.4; 1960a), Bharucha-Reid (1957; 1960, Chap. 4.4), Bailey (1964a, Chap. 12), and also Brambilla (1960) about continuous-time stochastic models. The important contributions of A. G. McKendrick (1926) to the theory of stochastic processes and their application to epidemiology are described in the Inaugural Address of J. O. Irwin (1963) to the Royal Statistical Society.

2. EPIDEMIOLOGIC CONCEPTS AND CLASSIFICATION OF MODELS

Before we can describe the models in detail, we have to define the most important concepts of epidemiology, since there are several different nomenclatures in medical textbooks. Even the notion of "epidemiology" has two meanings: (1) the study of the behaviour of infectious diseases in a population, or (2) the study of factors affecting the incidence of non-communicable disorders. We shall be concerned only with the first branch. For a general text on the subject the reader may refer to Taylor and Knowelden (1964).

Let us first illustrate the distinct epidemiological epochs involved in the infection of an individual A, and the spread of the infection to B (see Fig. 1; this is similar to Fig. 1 of Bartholomay, 1964):

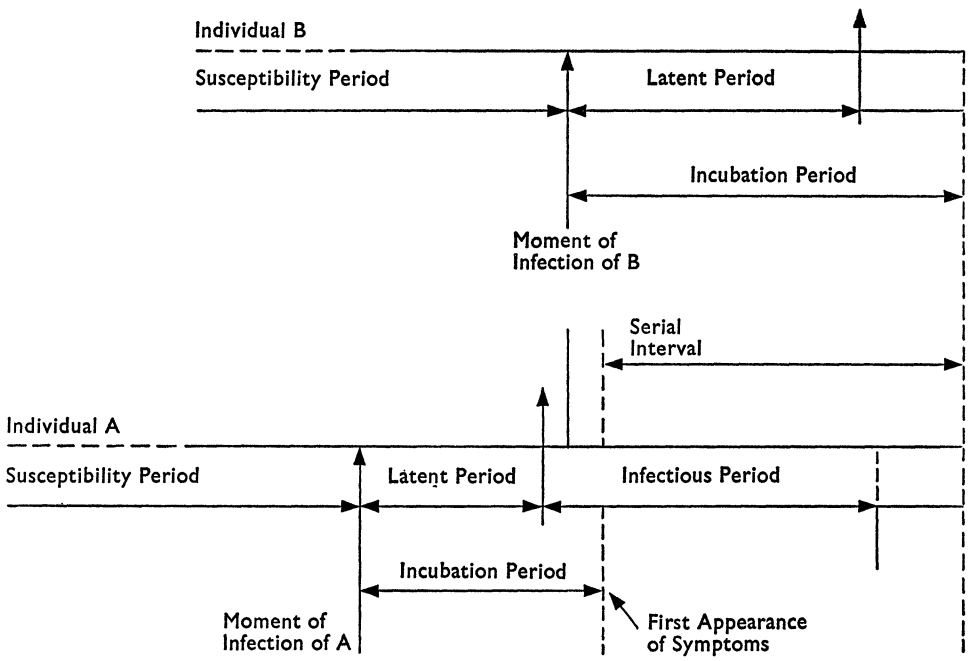

FIG. 1. Relations between epidemiological epochs.

An individual who may be infected by some contagious pathogenic agent is called susceptible. As soon as the causative organisms have been taken up by the body, a complicated process is initiated, during which the infectious material in general is not discharged. We call this time interval the latent period. This is immediately followed by the infectious period. The interval between the moment of infection and the appearance of symptoms is called the incubation period. Only after this time can preventive measures such as isolation be applied. The serial interval is the time between the observation of symptoms in individual A and the observation of symptoms in individual B, who is infected by A. This interval is the observable epidemiological unit. (Fig. 1 of Bartholomay, 1964, does not coincide with the definition of the serial interval given in the text.)

Almost all epidemic models have in common that they try to describe the number of infectives as a function of time, or of other pertinent variables such as the number

of susceptibles or of position. At least they predict the asymptotic qualitative behaviour. Deterministic models, which use a system of differential equations for the numbers considered, have been applied since 1906, starting with the work of Hamer. Since these models do not account for the considerable variations which occur in epidemic situations, stochastic models describing the change in time of certain distributions are much more appropriate for this kind of problem.

There are two main parts to mathematical epidemiology: the study of large-scale phenomena in large populations, and small-scale phenomena in households and schools. The necessity of the stochastic approach is obvious for the latter. The relevance of these models even for large-scale situations was pointed out by Bartlett (1964). One of the most important reasons is the possibility of extinction of a sub-population, e.g. of infectives. If their behaviour is described by a birth-and-death process with birth rate equal to death rate, then the deterministic analogue states that the number of infectives remains constant, whereas it follows from the stochastic theory that the population of infectives dies out with probability one. In addition, it is possible on the basis of the stochastic model to calculate the distribution of the time until extinction. In the above-mentioned paper Bartlett illustrates the inconvenience of the deterministic approach by a model for the spread of chickenpox and shingles. He indicates the possibilities of surmounting the greater mathematical difficulties involved in the stochastic treatment by use of Monte Carlo methods.

With a few exceptions we have to deal with temporally homogeneous Markov processes with a finite or countable state space. The most important classes of these are competition processes (see Reuter, 1961; Iglehart, 1964), branching processes (see Harris, 1963), birth-and-death processes (see Karlin, 1966, Chap. 7), and finite Markov chains (see Kemeny and Snell, 1960). We now characterize the different models as follows.

The *simple epidemic* describes the infection of a closed group of n susceptibles into which a infectives are introduced. We assume a comparatively mild disease such that there is no isolation of infectives; further it is supposed that the latent period is of length zero, and that the whole group of $n+a$ individuals is mixing homogeneously. This implies a growth rate of the infectives which is proportional to the product of the numbers of susceptibles and infectives. We shall discuss the recent contributions to this model by Bailey (1963; 1967a), Benayoun (1964), Kendall (1957), Mansfield and Hensley (1960), Severo (1967a), Störmer (1964), and Williams (1965a). For an extension of the simple deterministic epidemic to the case of two kinds of susceptibles, see Gart (1967) and Gart and De Vries (1966).

The *general epidemic* admits "removal" of infectives from circulation, either by isolation, death or immunity. In any case, the removals may not infect other susceptibles. Though this situation had been already studied by McKendrick (1926) and Bartlett (1949), a time-dependent solution appeared only recently (Gani, 1965; Siskind, 1965). There are also some new results about the total size of the general epidemic by Daniels (1967), Downton (1967a) and Cane (1966).

Individuals who are apparently healthy themselves, but are infective, are called *carriers*. Weiss (1965) proposed a model taking into account infection by carriers only. He considers a closed population of m susceptibles into which a number n of carriers is introduced. The latter are only detectable by the appearance of infected persons. The susceptibles may be infected until all carriers have been isolated or until the entire population has succumbed to the disease. A complete solution to this model was given by Dietz (1966) and Downton (1967b). See also Severo (1967b).

Dietz and Downton (1967) deal with the problem of immigration of carriers and susceptibles into the population. Downton (1967c) assumes that, after the initial introduction of carriers, no new carriers are introduced from outside, but that a proportion π of those susceptibles infected become carriers. For $\pi = 1$ and 0 this model coincides with the general epidemic and the model of Weiss, respectively. Pettigrew and Weiss (1967) consider a model involving two types of infectives in an infinite population of susceptibles such that they may apply branching processes.

The general epidemic assumes a latent period of length zero and an infectious period which has an exponential distribution. These assumptions are replaced by more realistic ones in a series of papers: Bailey (1957, Chap. 7) considers a normally distributed latent period and an infectious period of constant length. We shall not reproduce these results, but we shall discuss the approaches of Bharucha-Reid (1956, 1957), Bailey (1964b), Morgan (1964) and Bartoszyński (1967). In Section 6 we also shall report on the papers by Williams (1965b, c) and Puri (1966, 1967) about the distributions of incubation periods. See also Williams and Meynell (1967). They describe the development of the pathogenic organisms inside their hosts by a linear birth-and-death process. According to the assumptions of Williams, symptoms occur when the number of infecting particles reaches a fixed threshold. According to Puri, the number of infecting particles determines only the probability of response of the host. Thus their number at time of response is a random variable, the distribution of which is calculated by Puri (1967).

For the treatment of the spread of diseases which have a nearly constant latent period and relatively short infectious period, e.g. measles or mumps, the *chain-binomial models* of Reed and Frost are well adapted. For an excellent exposition of these models see Bailey (1957, Chap. 6). In mathematical terms, these are Markov chains with finite state space and discrete-time parameter. The unit of time is equal to the fixed length of the latent period. The period of infectiousness is concentrated into a single point, after which the individual becomes permanently immune. In Section 7 we shall deal with the extensions of Sugiyama (1961) and Elveback *et al.* (1964).

An essential assumption of all the above-mentioned models consists in the homogeneous mixing of the affected population. Especially for large populations this is unrealistic. Therefore one has to account for the fact that both susceptibles and infectives are geographically distributed and that infection between neighbours is more likely than between far-removed individuals. Section 8 considers five different approaches to this problem: (1) a deterministic model for the spread of epidemic waves by Kendall (1965); (2) a stochastic model of Neyman and Scott (1964) which admits movement of infectives; (3) percolation processes and their application to the spread of an infection of trees in an orchard; (4) Monte Carlo studies for the formation of plaques consisting of cells of a two-dimensional monolayer which are infected by virus particles (Schwöbel *et al.*, 1966); (5) Monte Carlo studies of Bailey (1967b) for the spread of an epidemic on a two-dimensional square lattice adopting a discrete-time model of chain-binomial type. The paper of Bailey (1967b) also contains a review of earlier deterministic models for the spatial spread of epidemics by Bartlett (1956) and Kendall (see discussion in Bartlett, 1957). See also Bailey (1967c).

A further complication is introduced into the mathematical theory of epidemics by diseases like malaria which involve an intermediate host (vector). Till recently, almost all approaches used deterministic models: Lotka (1923), D'Ancona (1954, Chap. 5), Macdonald (1957), Bailey (1957, pp. 29 *et seq.*; pp. 155 *et seq.*). Bartlett

(1964, 1966) calculates the probability of extinction of an epidemic under the assumption that the number of infectious hosts and of infectious vectors may be described by a branching process. (The number of susceptible hosts and vectors is kept constant.) Since there seem to be no further contributions to these types of epidemic models we shall confine ourselves to these marginal notes.

3. THE SIMPLE EPIDEMIC

3.1. *Solution of the Partial Differential Equation for the Probability Generating Function*

For simplicity we treat only the case of one initial infective ($a = 1$). Let $S(t)$ and $I(t)$ be the numbers of susceptibles and infectives, respectively, at time t. Then we always have the relation $S(t)+I(t) = n+1$, i.e. there are $S(t)$ susceptibles and $n+1-S(t)$ infectives. Since they are assumed to mix together homogeneously, the chance of one new infection in the whole group in $(t, t+dt)$ is $\beta(n-S+1)\,Sdt$. For the probabilities $p_s(t)$ that there are still s susceptibles uninfected at time t we introduce the probability generating function (p.g.f.)

$$P(x,t) = \sum_{s=0}^{n} p_s(t)\, x^s. \tag{3.1}$$

If we change the time scale to $\tau = \beta t$, then according to the well-known "random-variable" technique (see, e.g., Bailey, 1964a), equation (3.1) satisfies the following partial differential equation

$$\partial P/\partial\tau = (1-x)\{n(\partial P/\partial x) - x(\partial^2 P/\partial x^2)\}, \tag{3.2}$$

with the initial condition

$$P(x,0) = x^n.$$

Bailey (1963) developed an elegant method for the solution of (3.2) (see also Benayoun, 1964). In order to avoid multiple eigenvalues he replaces the integer n by $N = n+\epsilon$, deduces an explicit solution in terms of the hypergeometric function $F(a, b; c, x)$ (see Erdélyi, 1953), and finally obtains the desired result by letting $\epsilon \to 0$. This result had been previously derived by application of Laplace transforms. We do not reproduce the complicated formulae, which may be found also in Bailey's book (1957, pp. 40 *et seq.*). Severo (1967a) extends these results for arbitrary initial distributions.

Störmer (1964) gives asymptotic expressions for the distribution function of the number of infectives. If τ_i denotes the moment of infection of the ith infective ($i > a$), then we have

$$P\{I(t) \leqslant i\} = P\{\tau_{i+1} > t\} = 1 - P\{\tau_{i+1} \leqslant t\}.$$

For the distribution function of τ_{i+1} the following asymptotic relation is valid

$$P\left\{\beta\tau_{i+1} - \log n - \log\frac{i}{n-i} \leqslant \lambda\right\} = \sum_{k=0}^{a-1} \frac{(e^{-\lambda})^k}{k!} \exp(-e^{-\lambda}) + \epsilon(i, n-i),$$

where $\lim_{i,n-i\to\infty} \epsilon(i, n-i) = 0$. ($a$ is the initial number of infectives.) Störmer applies the simple epidemic to the description of the increase of consumption of consumer goods. (Bain, 1963, points out that a positively skew curve yields adequate description of the growth of demand for new commodities.)

Bailey (1967a) applies perturbation techniques in order to obtain manageable approximations for moderately large populations.

3.2. *The Epidemic Curve*

Notifications of new cases, as a function of time, constitute a histogram which is the statistical image of the epidemic curve. Mathematically this may be defined as $z(\tau) = dm(\tau)/d\tau$, where $m(\tau) = n+1-E\{S(\tau)\}$. The first explicit expressions for $m(\tau)$ was given by Haskey (1954). Bailey (1963) arrives at the same result by the method indicated above. It is interesting to compare $z(\tau)$ with the corresponding deterministic expression ($a = 1$)

$$\frac{n(n+1)^2 e^{(n+1)\tau}}{(n+e^{(n+1)\tau})^2}.$$

Figs. 2 and 3 demonstrate these differences for $n = 5$ and $n = 40$. The stochastic epidemic curves are drawn on the basis of data in the paper by Mansfield and Hensley

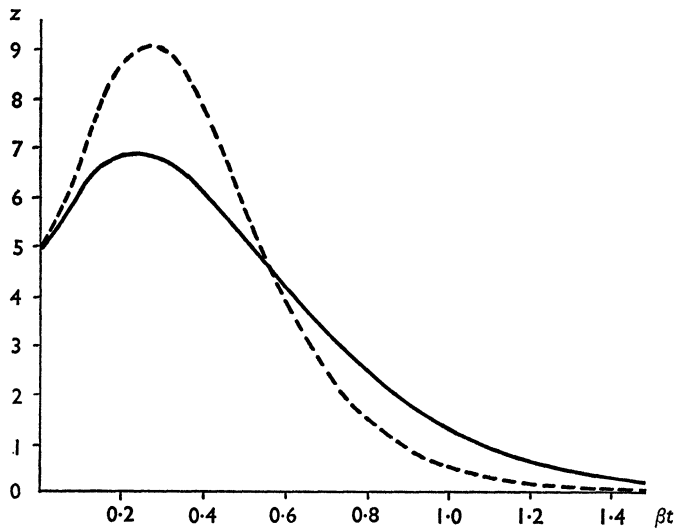

FIG. 2. Comparison of deterministic and stochastic epidemic curves for $n = 5$.
– – – – Deterministic curve; ——— stochastic curve.

(1960) where these are tabulated for $n = 5(1)40$. (Similar figures for $n = 10$ and 20 are given by Bailey, 1957, p. 21.) The differences between deterministic and stochastic solutions of the growth of the number of infectives has already been investigated by Feller (1939). Indeed, the number of infectives increases according to a logistic process with zero death rate. Feller showed that the deterministic solution always exceeds the corresponding mean value.

A more detailed comparison of the deterministic and stochastic solutions has been performed by Williams (1965a). He interprets the properly normalized epidemic curve as the frequency function of the time of occurrence of a new case. Then the mean of the deterministic curve equals

$$n^{-1}\log\{1+(n/a)\}$$

and the stochastic mean turns out to be

$$n^{-1}\sum_{r=a}^{n+a-1} r^{-1}.$$

For n and a tending to infinity the difference vanishes, but it is noticeable for small values of a. By comparing the excesses $\gamma_2 = \kappa_4/\kappa_2^2$ of both curves for $a = 1$ and $n \to \infty$, Williams shows that γ_2 of the stochastic curve is two-thirds of the corresponding deterministic value (= 1·2 exactly). (Cf. Figs. 2 and 3.)

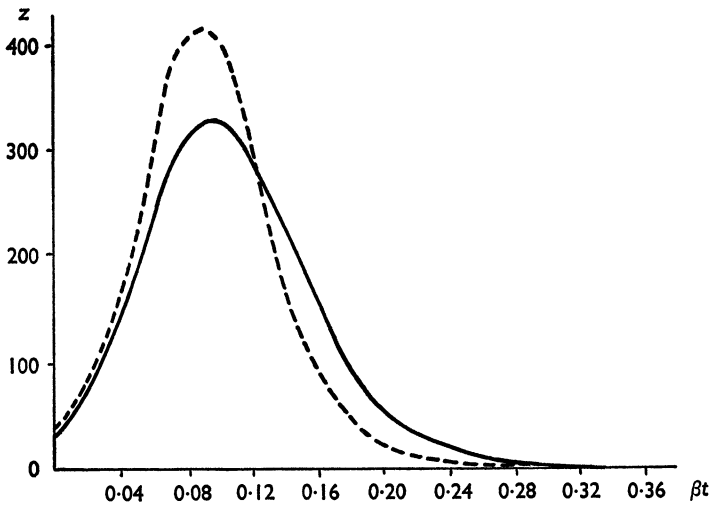

FIG. 3. Comparison of deterministic and stochastic epidemic curves for $n = 40$.
– – – – – Deterministic curve; ——— stochastic curve.

Kendall (1957) gives asymptotic expressions for the distribution of the duration time of the simple epidemic in terms of modified Bessel functions. For the results, see Bailey (1957, p. 49). For $n \to \infty$, the mean duration time tends to zero—a consequence of homogeneous mixing!

4. THE GENERAL EPIDEMIC

4.1. *Solution of the Partial Differential Equation for the Probability Generating Function*

Following the assumptions for the general epidemic made in Section 2, we consider the following transition rates of the stochastic model (m and n denote the numbers of susceptibles and infectives, respectively):

$$\begin{aligned} &(m,n) \to (m-1, n+1) && mn \\ &(m,n) \to (m, n-1) && \rho n \\ &(m,n) \to (m,n) && -(m+\rho)n. \end{aligned}$$

The time scale is chosen such that the infection rate equals 1, while ρ is called relative removal rate. We consider a homogeneously mixing population of s susceptibles and a infectives. For the probabilities $p_{m,n}(\tau) = P\{m$ susceptibles, n infectives at time $\tau | s$ susceptibles, a infectives at time $0\}$ we introduce the p.g.f.

$$P(x,y,\tau) = \sum_{m=0}^{s} \sum_{n=0}^{s+a-m} p_{m,n}(\tau)\, x^m y^n. \tag{4.1}$$

This satisfies the partial differential equation

$$\partial P/\partial\tau = y(y-x)(\partial^2 P/\partial x\,\partial y)+\rho(1-y)(\partial P/\partial y), \tag{4.2}$$

with the initial condition

$$P(x,y,0) = x^s y^a.$$

Recently, Gani (1965, 1967) and Siskind (1965) solved (4.2) independently at about the same time. In principle, both apply the same method. Gani succeeds in deriving explicit expressions for the Laplace transforms of the individual probabilities, while Siskind obtains an explicit expression for the p.g.f. (4.1) which is not reproduced here because of its length.

We shall confine ourselves to the characterization of their common method. The idea consists in the reduction of a complicated second-order partial differential equation to a system of first-order partial differential equations, which is further reduced by Gani to a system of ordinary differential equations by the application of Laplace transforms. This reduction is achieved by the following substitution:

$$P(x,y,\tau) = \sum_{m=0}^{s} x^m f_m(y,\tau),$$

where

$$f_m(y,\tau) = \sum_{n=0}^{s+a-m} p_{m,n}(\tau)\,y^n.$$

The system for the f_m is obtained by equating the coefficients:

$$\partial f_m/\partial\tau = y^2(m+1)(\partial f_{m+1}/\partial y)-\{(m+\rho)y-\rho\}(\partial f_m/\partial y) \quad (r=0,1,\ldots,s-1),$$

$$\partial f_s/\partial\tau = -\{(s+\rho)y-\rho\}(\partial f_s/\partial y),$$

with the initial condition

$$f_s(y,0) = y^a, \quad f_m(y,0) = 0 \quad (m=0,1,\ldots,s-1).$$

This system can be solved fairly straightforwardly. Severo (1967a) also considers arbitrary initial distributions for this model.

4.2. *The Total Size of the General Epidemic*

The total size of the general epidemic is $s-S(\infty)$, i.e. the number of all cases occurring during the course of the epidemic. If this model is treated deterministically, there results an asymptotic relation between this number and the ratio of the initial number of susceptibles s to the relative removal rate: $s<\rho$ implies no epidemic, but $s>\rho$ means that the initial number of susceptibles $s=\rho+\delta$ is reduced to $\rho-\delta$. This is the well-known threshold theorem of Kermack and McKendrick (1927). A more detailed asymptotic analysis yields $se^{-s/\rho}$ as the final value of the number of susceptibles. (See also Marchand, 1956.)

Bailey (1953) obtained a stochastic analogue of this theorem: For $s<\rho$ the corresponding distribution of the total size is J-shaped, but for $s>\rho$ it is U-shaped, i.e. there are only minor or major outbreaks, and only a few of intermediate size. (For a graphical illustration of these distributions, see Spicer and Lipton, 1958.) Whittle (1955) and Kendall (1956) approximated the general epidemic by birth-and-death processes which are extinguished with probability one in the case $s\leqslant\rho$ and with probability $(\rho/s)^a$ in the case $s>\rho$.

An approximate treatment of the total size of the general epidemic has been given by Daniels (1967), who obtained quite good coincidence with the individual probabilities. He showed by analysing the embedded random walk that

$$p(r\,|\,s,a,\rho) = \binom{s}{r}\left(\frac{\rho}{\rho+r}\right)^{s+a-r} p(0\,|\,s-r,a,\rho+r), \tag{4.3}$$

where $p(r\,|\,s,a,\rho)$ is the probability that the epidemic terminates with r susceptibles remaining uninfected given that the initial numbers of susceptibles and infectives equal s and a, respectively, and that the relative removal rate is equal to ρ. Relation (4.3) implies that it is only necessary to know these probabilities for $r = 0$. Downton (1967a) derives (4.3) by a simple combinatorial argument.

Of the various asymptotic results of Daniels we only mention the following one: For ρ large and s much larger, the distribution of the number remaining uninfected in a large epidemic has approximately the Poisson form with the deterministic mean $se^{-s/\rho}$. Cane (1966) only considers first and second moments of the general epidemic and arrives by heuristic arguments at the same result: "In the case discussed by Daniels, we have a large population exposed to infection, a small probability that any particular member of it will remain uninfected, and the fate of one individual is virtually independent of the fate of another, so that we should expect the distribution of the number uninfected to be approximately Poisson."

4.3. *Estimation of Parameters*

So far, we have only dealt with one parameter in the general epidemic, but in this section we shall consider the infection rate λ and the removal rate μ separately ($\rho = \mu/\lambda$). In order to estimate λ and ρ, Bailey (1957) proposed the following method.

Consider households consisting initially of one susceptible and one infective. It is assumed that no infection from outside occurs. Let A denote the event that there is no further case of infection, and B denote the complement of A. Then the probabilities of these events are

$$P(A) = \mu/(\lambda+\mu); \quad P(B) = \lambda/(\lambda+\mu).$$

The distribution function of the moment of removal is $F(t) = 1-e^{-\mu t}$. By classifying the households according to the occurrence of A or B and the observed removal times, Bailey calculated maximum-likelihood estimators $\hat{\lambda}$ and $\hat{\mu}$:

$$\hat{\mu} = N_2/L; \quad \hat{\lambda} = N_2^2/N_1 L,$$

where N_1 and N_2 are the numbers of households with exactly one or two cases, respectively, and L denotes the sum of all case-to-case intervals.

Ohlsen (1964) notes that this method can be improved, since the exponential distribution admits a very long time between the first and second removals, and therefore it is in principle impossible to say that a one-removal household is an example of event A. Furthermore, the observation of a very long removal time after the first case suggests the conjecture of a second infection from outside rather than from contact with the initial infective. Ohlsen proposes the introduction of a fixed interval of observation of length T: Depending on the occurrence of a second case during

this interval the households are associated with classes A or B. Ohlsen gives the maximum-likelihood estimators $\hat{\lambda}$ and $\hat{\mu}$ for this modified method as

$$\hat{\lambda} = \frac{N_2 \hat{\mu}}{N_1 - (N_1 + N_2) e^{-\hat{\mu} T}},$$

$$\hat{\mu}\{N_2 T e^{-\hat{\mu} T} + (1 - e^{-\hat{\mu} T}) T\} = N_2 (1 - e^{-\hat{\mu} T}),$$

where $\hat{\mu}$ is the largest root of the last equation. She notes that these estimators tend to those of Bailey for $T \to \infty$.

For small values of T, however, there are striking differences, which are investigated in detail by Morgan (1965). In addition, Morgan admits infection from outside and calculates an estimator for its probability.

4.4. *Extensions of the General Epidemic*

The model of the general epidemic has been extended in several directions. For instance, Sakino and Hayashi (1959) and Sakino (1962b) admit the direct transition of a susceptible into the removal state, without intervening infection. (From the medical standpoint this assumption seems unrealistic.) Another extension is concerned with immigration of susceptibles or infectives into a population. The immigration of new susceptibles and infectives gives rise to recurrent epidemics. Attempts have been made to describe these phenomena by deterministic models, but the solutions turned out to be damped oscillations, in contrast to the observations. Therefore a stochastic analysis was undertaken by Bartlett (1956, 1957, 1960a, b, c, 1961). (See also Black, 1966; Goffman, 1965; Seiden, 1957; Taylor, 1958). Bartlett assumes the following transition rates:

$$\begin{aligned}
&(m,n) \to (m-1, n+1) && \lambda mn \\
&(m,n) \to (m, n-1) && \mu n \\
&(m,n) \to (m+1, n) && \nu \\
&(m,n) \to (m, n+1) && \epsilon \\
&(m,n) \to (m, n) && -(\lambda mn + \mu n + \nu + \epsilon),
\end{aligned}$$

where m and n are the numbers of susceptibles and infectives, respectively.

Ridler-Rowe (1967) obtained the following results, assuming these rates:

(1) By application of a general theorem of Reuter (1957, 1961), Ridler-Rowe proves an asymptotic formula for the expectation of the time $T(s, a)$ until the extinction of the infectives for the first time:

$$E\{T(s,a)\} \sim \mu^{-1} \log(s+a) \quad (s+a \to \infty, a > 0).$$

As above, s and a denote the initial numbers of susceptibles and infectives.

(2) The process is "non-dissipative", i.e. the sum of all individual probabilities of the limiting distribution equals one. In other words, the process tends to infinity for $t \to \infty$ only with probability zero.

(3) If $\nu \to 0+$, i.e. if the immigration rate of susceptibles tends to zero, then the number of infectives is distributed asymptotically like a Poisson variable with mean ϵ/μ, where the number of susceptibles equals zero.

A method of solution of the corresponding partial differential equation for the p.g.f. of the above model has been indicated by Gani (1966). But the expressions derived are of such unwieldy length that some judicious approximations may give

more insight into the properties of the model. For examples of these the reader is referred to the papers by Bartlett mentioned above.

Finally, we mention the papers of Sakino (1959, 1962c, 1963), taking into account the age distribution of the infectives.

5. CARRIER-BORNE EPIDEMICS

5.1. *Time-dependent Transition Rates*

The model of Weiss (1965) assumes the following transition rates:

$$\begin{aligned} (m,n) &\to (m-1,n) && \alpha mn \\ (m,n) &\to (m,n-1) && \beta n \\ (m,n) &\to (m,n) && -(\alpha m+\beta)n. \end{aligned}$$

Let (M,N) be the initial state. Here the first and second coordinates indicate the numbers of susceptibles and carriers, respectively. The numbers of carriers and susceptibles decrease according to stochastic death processes. Whereas the carriers continue to be eliminated as long as their number is non-zero, the number of susceptibles decreases only as long as there are still carriers within the population. Unlike Weiss, we admit the dependence of α and β on time in order to account for the change of environmental conditions in the spread of the epidemic.

It is, for example, reasonable to assume an increasing function of β because the isolation measures of public health authorities become more efficient during the course of the epidemic.

According to these assumptions we have the following partial differential equation for the p.g.f.

$$P(x,y,t) = \sum_{m=0}^{M} \sum_{n=0}^{N} p_{m,n}(t)\, x^m y^n.$$

$$\partial P/\partial t = \alpha(t)(1-x)y(\partial^2 P/\partial x\,\partial y) + \beta(t)(1-y)(\partial P/\partial y) \tag{5.1}$$

with the initial condition $P(x,y,0) = x^M y^N$. If we introduce the functions $f_r(y,t)$ by the following relation

$$P(x,y,t) = \sum_{r=0}^{M} \binom{M}{r} (x-1)^r f_r(y,t),$$

and substitute this in (5.1), we then obtain

$$\sum \binom{M}{r} (x-1)^r [(\partial f_r/\partial t) + \{(r\alpha+\beta)y - \beta\}(\partial f_r/\partial y)] = 0. \tag{5.2}$$

Since the polynomials

$$P_r(x) = \binom{M}{r} (x-1)^r$$

are linearly independent, (5.2) yields a simple first-order partial differential equation for $f_r(y,t)$, the solution of which is

$$f_r(y,t) = \left\{ y e^{-F_r(t)} + \int_0^t \beta(\tau)\, e^{-F_r(\tau)}\, d\tau \right\}^N,$$

where

$$F_r(t) = \int_0^t \{r\alpha(\tau) + \beta(\tau)\}\, d\tau.$$

5.2. *The Number of Susceptibles*

For the expectation of the number of susceptibles $M(t)$ we easily obtain

$$E\{M(t)\} = M\left(e^{-F_1(t)} + \int_0^t \beta(\tau)\, e^{-F_1(\tau)}\, d\tau\right)^N.$$

In the case $\alpha(t) \equiv \alpha$, we may give a closed expression for $\lim_{t\to\infty} E\{M(t)\} = M_\infty$:

$$M_\infty = M\left[\int_0^\infty \beta(\tau) \exp\left\{-\alpha\tau + \int_0^\tau \beta(\xi)\, d\xi\right\} d\tau\right]^N = M\{1-\alpha\mathscr{L}(\alpha)\}^N,$$

where $\mathscr{L}(.)$ is the Laplace transform of

$$\exp\left\{-\int_0^t \beta(\tau)\, d\tau\right\}.$$

Suppose, for example, that we set $\beta(t) = \beta + \gamma t$, we find

$$M_\infty = M\big(1-\alpha(2\pi/\gamma)^{\frac{1}{2}} \exp\{(\alpha+\beta)^2/2\gamma\}\,[1-\phi\{(\alpha+\beta)/\gamma^{\frac{1}{2}}\}]\big)^N,$$

where

$$\phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x \exp(-\tfrac{1}{2}\xi^2)\, d\xi.$$

For $\beta(t) = \beta - \{b/(1+t)\}$ we have

$$M_\infty = M\{1-\alpha(\alpha+\beta)^{-(b+1)}\, e^{\alpha+\beta}\, \Gamma(b+1, \alpha+\beta)\}^N,$$

where

$$\Gamma(a, x) = \int_x^\infty e^{-t}\, t^{a-1}\, dt$$

is the incomplete gamma function. If we set $b = 0$, the result of Weiss (1965) is obtained:

$$M_\infty = M\{\beta/(\alpha+\beta)\}^N.$$

From the corresponding differential equations of the deterministic analogue we obtain for the number of susceptibles $M_d(t)$:

$$M_d(t) = M \exp\left\{-N \int_0^t \alpha(\tau)\, e^{-B(\tau)}\, d\tau\right\},$$

where

$$B(t) = \int_0^t \beta(\tau)\, d\tau.$$

It turns out that the stochastic model yields a greater value than $M_d(t)$.

5.3. *Duration of the Epidemic*

The epidemic may terminate either by elimination of all carriers or by the infection of all susceptibles before carriers are eliminated. In the case $\alpha(t) \equiv \alpha$ and $\beta(t) \equiv \beta$, we obtain for the distribution function $F(t)$ of the duration of the epidemic

$$F(t) = (1-e^{-\beta t})^N + \sum_{m=0}^{M} (-1)^m \binom{M}{m} \left\{\left(\frac{\alpha m e^{-(\alpha m+\beta)t} + \beta}{\alpha m + \beta}\right)^N - \left(\frac{\beta(1-e^{-(\alpha m+\beta)t})}{\alpha m+\beta}\right)^N\right\}.$$

The following integral yields an expression for the mean time of duration

$$\int_0^\infty \{1-F(t)\}\,dt = \sum_{n=1}^{N} \frac{1}{\beta n}\left\{(-1)^{n-1}\binom{N}{n} - \sum_{m=0}^{M}(-1)^m \binom{M}{m}\left(\frac{\beta}{\alpha m+\beta}\right)\right\}^{N-n+1}.$$

Downton (1967b) indicates how this may be simplified by the introduction of polygamma functions.

6. Distributions of Latent, Incubation, and Infectious Periods

6.1. *Age-dependent Branching Processes*

Bellman and Harris (1948) developed a theory of branching processes which admits an arbitarily distributed generation time of the reproducing particles. For details the reader is referred to Chapter 6 in Harris (1963). Bharucha-Reid (1956) introduced this model into the theory of epidemics in order to account for latent periods with a general distribution G. The inconvenience of this model consists in neglecting the number of susceptibles, since only the number of infectives is described by a one-dimensional branching process. After completion of the latent period an infective may infect n new individuals with probability q_n ($\sum_{n=0}^{\infty} q_n = 1$), whereupon he will be isolated at once. Bharucha-Reid (1956) considers only the case $G(t) = 1-e^{-bt}$ and $q_n = 0$ for $n \geqslant 3$ in detail. In addition he indicates a sequential method for the comparison of two special branching processes (see also Bharucha-Reid, 1958).

6.2. *Latent Periods and Infectious Periods with General χ^2-distributions*

Bailey (1964b) and Morgan (1964) consider latent periods and infectious periods which are distributed as χ^2 with an even number of degrees of freedom. In order to retain the Markov property of the stochastic model they introduce imaginary states, as was done for the first time by Kendall (1948) in describing the growth of a population. One assumes that an individual has to pass through k stages after infection during the latent period and through l stages during the infectious period, where the time spent in each stage is exponentially distributed, i.e. like a χ_2^2 variable. For the entire latent and infectious periods one obtains a χ_{2k}^2-distribution and a χ_{2l}^2-distribution, respectively, since the individual times are independent. The above-mentioned papers contain no complete solutions.

6.3. *Latent Periods and Infectious Periods with Discrete Distributions*

Another method of dealing with general latent periods and infectious periods is due to Bartoszyński (1967). Like Bharucha-Reid (1956) he applies a one-dimensional branching process to describe the number of infectives; specifically he uses the discrete time Galton–Watson process (see Harris, 1963, Chap. 1). The p.g.f. for the number of cases which are infected by one infective is derived under the following assumptions:

(1) Every infected individual passes through a latent period of length X and through an infectious period of length Y with the common distribution

$$p_{ij} = P\{X = i,\ Y = j\}$$

and the generating function

$$F(x, y) = \sum_{i,j} p_{ij}\, x^i y^j,$$

where the times are measured in discrete units (e.g. days).

(2) On each day during the latent and infectious periods an infected individual may be discovered and isolated with probability $1-\alpha$ $(0<\alpha\leqslant 1)$ and $1-\beta$ $(0<\beta\leqslant 1)$, respectively.

(3) The random number of contacts of every infected individual with susceptibles during a day has the p.g.f. $R(y) = \sum_k r_k y^k$ with $R'(1)<\infty$.

(4) The probability of a "successful" contact is γ $(0<\gamma\leqslant 1)$.

(5) The events occurring to an individual on one day are independent of the events which occur to him or other individuals on previous days, and are independent of the events which occur to other members of the population on the same day.

Under these assumptions Bartoszyński arrives at the following p.g.f. for the number of cases infected by an infective:

$$P(y) = 1-F(\alpha,1)+\frac{(1-\beta)F(\alpha,1)+\beta\{1-R(\gamma y+1-\gamma)\}F\{\alpha,\beta R(\gamma y+1-\gamma)\}}{1-\beta R(\gamma y+1-\gamma)}.$$

It is possible to investigate the conditions for which there is a positive probability for an epidemic, i.e. when the number of infectives tends to infinity. For this it is necessary that $P'(1)>1$. Thus, if the latent period has constant length M and the infectious period has mean length D, then there is no epidemic for either

$$\gamma\alpha^M\leqslant\{R'(1)\,D\}^{-1} \quad (\beta \text{ arbitrary}) \qquad \text{or} \qquad \beta\leqslant\{1+\gamma R'(1)\,\alpha^M\}^{-1}.$$

Even for infinite D it will be possible to prevent an epidemic by a proper choice of β (e.g. by means of periodic check-ups).

6.4. *Distribution of Incubation Periods*

Sartwell (1950) fitted histograms of incubation periods by the log-normal distribution and obtained good coincidence. A Bayesian method for the estimation of the parameters in connection with point-source epidemics is given by Hill (1963).

Williams (1965b, c) tries to explain the positive skewness of the empirical distributions by means of stochastic models for the development of the pathogenic organisms *in vivo*. Meynell and Meynell (1958) report experimental results tending to support the hypothesis of a fixed threshold at which the response of the host occurs. Under the assumption that the development of the micro-organisms is described by a linear birth-and-death process, Williams calculated the asymptotic distribution of the time until reaching the threshold N for large values of N (see also Saaty, 1961a; b, p. 117). This model explains the long tails to the right of observed distributions of incubation periods: a brief response time demands an improbable path to the threshold, with few or no reverses; on the other hand, there are many possible paths by which the process can arrive late at the threshold. Similar models for incubation times involving a fixed threshold have been considered by Gart (1965).

Puri (1966, 1967) abandons the hypothesis of a fixed threshold. The occurrence of the response of the host is a random event, the probability of which depends on the number of existing living particles $N(t)$ as well as on the integral $\int_0^t N(u)\,du$. The latter term is considered as a measure of the amount of toxin produced by the live particles during the interval $(0, t)$. The model of Puri does not take into account a defence mechanism of the host. This point should be regarded in further approaches to these phenomena.

7. CHAIN-BINOMIAL MODELS

7.1. *The Model of Sugiyama*

Let $S(t)$ be the number of susceptibles just prior to time t, and let $I(t)$ be the number of infected individuals just prior to time t who will become infectious at that instant. If $p = 1-q$ is the probability of a "successful" contact between any two members of the group at time t, then the probability of no contact of a given susceptible with any of the $I(t)$ infectives is $q^{I(t)}$. Accordingly the probability of a contact with at least one infective is $1-q^{I(t)}$. The conditional probability of $I(t+1)$ new infections is given by the binomial distribution

$$P\{I(t+1)\,|\,S(t), I(t)\} = \binom{S(t)}{I(t+1)} (1-q^{I(t)})^{I(t+1)} q^{I(t)\,S(t+1)},$$

where $S(t) = S(t+1)+I(t+1)$. The process develops according to a chain of binomial distributions; this is the reason for the name of these models.

All chain-binomial models so far assumed a certain initial number of infectives in a small group of individuals and excluded the possibility of an outdoor infection, once the primary cases occurred. Sugiyama (1961) abandons this unrealistic assumption and introduces two parameters into his model: p_0 is the probability that a susceptible contracts a certain disease by *outdoor* infection during a unit period of time; p_i is the corresponding probability of infection by *one infectious case within* the same household who had been infected during the preceding time period. As an example, Sugiyama deals with households of three susceptibles and gives a list of the probabilities for all 20 possible realizations of this process if this is considered only at three moments of time T, $2T$, $3T$. We shall confine ourselves to the realization (1–1–0), i.e. the numbers of infected cases at $T, 2T, 3T$ equal $1, 1, 0$ respectively, the probability of which is clearly

$$3p_0(1-p_0)^2\,2\{1-(1-p_0)(1-p_i)\}\{(1-p_0)(1-p_i)\}(1-p_0)(1-p_i).$$

On the basis of the observed numbers of households with the different types of realizations, the maximum-likelihood estimators for p_0 and p_i are calculated. The comparison with data of 42 households during the A-Asia 57 influenza epidemic in Osaka yields good coincidence ($\chi_1^2 = 2{\cdot}1$).

7.2. *The Model of Elveback et al.*

Elveback *et al.* (1964) discuss the question whether interference of two competing viral agents A and B would be detectable by field observations of the total number of cases after completion of the epidemics. They use Monte Carlo methods in order to avoid complicated calculations. (See also Horiuchi and Sugiyama, 1957, for further Monte Carlo studies on chain-binomial models.) Whereas the usual chain-binomial models distinguish only three states, namely susceptible, infective and immune, we now have the eight following ones:

1. Susceptible to A and B.
2. Susceptible to A, immune to B.
3. Susceptible to B, immune to A.
4. Immune to A and B.
5. Infective to A, temporarily insusceptible to B.
6. Infective to A, permanently immune to B.
7. Infective to B, temporarily insusceptible to A.
8. Infective to B, permanently immune to A.

The possible transitions during the next interval are illustrated in Fig. 4.

The assumptions of the above-mentioned eight states imply complete interference of A and B, since the presence of an infection with A reduces the probability of infection with B to zero and vice versa. Under these optimum conditions for the demonstration of the existence of interference, however, it turned out that there was

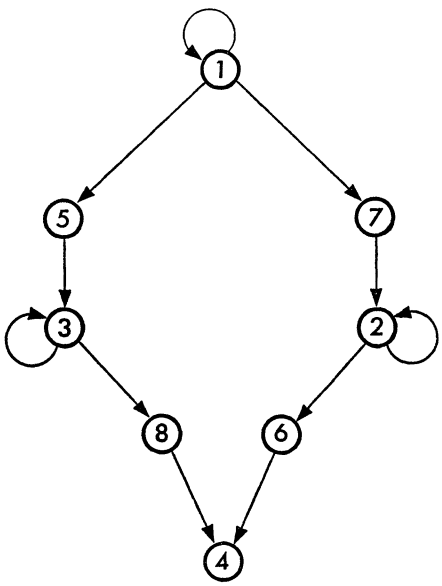

FIG. 4. State space and possible transitions (→) of the model of Elveback *et al.*

no significant difference in the distribution of the total number of A cases for B present or not. Thus, study of competition by means of household distributions is not useful. (Goffman, 1966a, investigates conditions for the simultaneous occurrence of several infectious diseases.)

8. SPATIAL SPREAD OF EPIDEMICS

8.1. *A Deterministic Model of D. G. Kendall*

Kendall (1965) considers a one-dimensional population of uniform density σ and states conditions for the existence of waves of infection travelling along this linear community. Let $s(x, t)\,\sigma$ and $i(x, t)\,\sigma$ be the local density of susceptibles and infectives, respectively, at time t and position x. Then these densities are assumed to satisfy the following differential equations:

$$\partial s/\partial t = -s\{i+k(\partial^2 i/\partial x^2)\}; \quad \partial i/\partial t = s\{i+k(\partial^2 i/\partial x^2)\}-\rho/\sigma^i.$$

We note that the rate of occurrence of infections at a point is proportional to the density of susceptibles and to a quantity which approximates a local average of i, where $k>0$; ρ is the removal rate of infectives. Kendall looks for travelling waves by putting

$$s(x, t) = s(x-ct), \quad i(x, t) = i(x-ct),$$

where c (>0) is the velocity of the waves. He obtains two important results:

(1) Unless the population density σ exceeds the threshold value ρ, there are no waves.
(2) No waves are possible with a speed smaller than

$$c_{\min} = 2\{k(\sigma-\rho)/\sigma\}^{\frac{1}{2}}$$

but waves are possible with every velocity greater than or equal to $c_{\min}$.
This approach could be extended for a two-dimensional population.

8.2. *A Stochastic Model of Neyman and Scott*

The model of Neyman and Scott (1964) accounts for the spatial distribution of susceptibles and infectives and for the movement of infectives within the habitat H. The number of susceptibles, in particular its decrease, is neglected. The model may be characterized as a position-dependent branching process with discrete time parameter, where the unit of time corresponds to the latent period.

The phenomenon of dispersal of infectives is described by the density $f(v|u)$. Here v is the position at which an individual infected at u becomes infectious. Hence

$$p(R|u) = \int_R f(v|u)\,dv$$

is the probability that an individual infected at u becomes infectious somewhere within the sub-region R of the habitat. $I(v)$ denotes the number of susceptibles which will be infected at v if there is an infective, and $g(y|v)$ denotes its p.g.f.

In addition to these assumptions Neyman and Scott admit immigration of infectives according to the position-dependent Poisson rate $\zeta(v) \geqslant 0$. Thus the p.g.f. of the number of immigrated infectives $Z(R)$ in the region R during one unit of time is

$$F\{Z(R|x)\} = \exp\left\{-(1-x)\int_R \zeta(v)\,dv\right\}.$$

This model reduces to a branching process with a finite number of types if the habitat is assumed to be composed of a finite number of points. The theory of these processes is presented in Chapter 2 of Harris (1963). Bartoszyński (1967) also has generalized the model of Section 6.3 for several disjoint regions in the plane.

The following problems are dealt with by Neyman and Scott:

(1) Let $R_1, R_2, \ldots, R_k$ be arbitrary, disjoint regions and let $N_{1i}(u)$ be the number of individuals infected at u which become infectious at the beginning of the next time unit in the region R_i $(i = 1, \ldots, k)$. The p.g.f. of the vector $N_1(u)$ with components $N_{1i}(u)$ has the form

$$G_1(\mathbf{y}|u) = g\left\{1 - \sum_{i=1}^{k}(1-y_i)p(R_i|u)\middle|u\right\}.$$

(2) From this follows immediately the form of the p.g.f. for the number of infectives infected by a single individual who has himself been infected at u:

$$h_1(\mathbf{y}|u) = \int_H G_1(\mathbf{y}|v)f(v|u)\,dv.$$

(3) Similarly one obtains the p.g.f. of the number of infectives in the nth generation of an epidemic.

(4) Neyman and Scott give sufficient conditions under which an epidemic started by one individual is destined to become extinct with probability one. Bühler (1966) gives a weaker, but still only sufficient, condition.

(5) Further, the course of an epidemic which is continuously generated by immigrants is investigated.

(6) Finally, they evaluate the influence of an immunization campaign on the total size of epidemic. Under certain conditions they arrive at the optimistic result of a reduction of the expected size to a value less than $(1-\theta)/\theta$ had a random proportion θ been immunized. (Two special distributions of the total size of a branching process are presented by Neyman, 1965.)

Bartoszyński *et al.* (1965) investigate conditions under which, with probability one, an epidemic either becomes extinct or grows indefinitely (since no restrictions are introduced on the size of the population in the Neyman–Scott model).

8.3. *Percolation Processes*

For a review of the theory and applications of percolation processes, the reader is referred to Frisch and Hammersley (1963). We only cite two characteristic examples for the theory from Hammersley and Welsh (1965):

(1) Let us consider a large orchard in which the trees are planted at the vertices of a square lattice. The distance between trees is such that an infected tree may only infect its four nearest neighbours, each with probability p. Then one is interested in the distribution of the number of infected trees.

(2) Now it is supposed that a given tree, once infected, can only infect a given neighbour after a random time interval T. First-passage percolation theory considers the time at which infection first spreads outside a given region.

For the case of exponentially distributed T the reader should consult the papers by Morgan and Welsh (1965) and Hammersley (1966).

8.4. *A Model for the Formation of Plaques*

Schwöbel *et al.* (1966) identify the cells of a two-dimensional monolayer with regular hexagons. The number of infective units produced by a single cell just infected is a Poisson-distributed random variable. The infective units are distributed among the six neighbouring cells with equal probability. One infective unit is sufficient to infect a cell; a further infective unit hitting the cells remains without effect. The process develops in discrete time units corresponding to the generation times of the infective agents. Schwöbel *et al.* apply Monte Carlo methods in order to estimate the probability of extinction as a function of the Poisson parameter λ. They have found a linear relationship of the radius of model plaques with time, the speed of this process depending on λ.

8.5. *Two-dimensional Chain-binomial Models*

Bailey (1967b) reports on Monte Carlo studies of epidemics which spread according to chain-binomial type models (Section 7) through a population of susceptibles which are located at the points of a square lattice. The eight nearest neighbours may be infected (unless already infected) by an infectious point. Bailey considers simple epidemics without recovery and general epidemics with infectives undergoing removal. The Monte Carlo results are demonstrated by numerous figures.

9. RUMOURS

9.1. *Epidemic Models and the Spread of Information*

Most of the above-mentioned epidemic models have a very general structure which is applicable also to the description of other phenomena. Thus the logistic

process, which is equivalent to the simple epidemic of Section 3, has been used by Kendall (1957) in order to predict the spread of a message through a finite and closed population.

For the rumour interpretation of the model the susceptibles are identified with those not having heard the rumour and the infectious cases correspond to those who are actively spreading the rumour. Goffman and Newill (1965) interpret epidemic models involving a vector as models for the spread of ideas by publications. Goffman (1966b) describes the development of mast cell research by the general epidemic. There is also a series of papers about a deterministic theory of rumour spread by Rapoport and Rebhun (1952), Rapoport (1953a, b), Landau and Rapoport (1953), Landahl (1953) and Daley (1967). Daley and Kendall (1964, 1965) point out that the spread of rumours is governed by other laws than the spread of epidemics. We shall describe their model in the next section.

9.2. *The Model of Daley and Kendall*

The model of Daley and Kendall for rumour spread has some properties in common with the general epidemic treated in Section 4. Both models suppose a closed population of N susceptibles or ignorants (class A), with one initial infective or spreader (class B), and a contagion effect transferring persons from class A to class B at a rate proportional to XY, where X and Y are the numbers (random) of persons in class A and B, respectively. In the general epidemic, persons formerly infectives will be removed (come into class C) at a rate proportional to Y by death, isolation or recovery. At this point the models differ: an active spreader of the rumour is removed whenever he encounters somebody who has heard it before. If he meets a member also of class B two persons are transferred to class C, but if he meets a member of class C only one person is transferred to C. The corresponding transition rates are as follows (Z is the number of stiflers):

$$\begin{aligned}
(X, Y, Z) &\to (X-1, Y+1, Z) && XY \\
(X, Y, Z) &\to (X, Y-2, Z+2) && \tfrac{1}{2}Y(Y-1) \\
(X, Y, Z) &\to (X, Y-1, Z+1) && YZ \\
(X, Y, Z) &\to (X, Y, Z) && -Y(X+\tfrac{1}{2}(Y-1)+Z).
\end{aligned}$$

Daley and Kendall apply several approximative methods to this stochastic model. Their main interest is directed to the evaluation of the final proportion f of persons ultimately hearing the rumour. The deterministic analogue yields $f \approx 0{\cdot}80$. The method of the embedded random walk gives nearly the same result together with values for the variance σ_N^2 of the number of persons hearing the rumour. Thirdly, a new method, which they call the *Principle of the Diffusion of Arbitrary Constants*, requiring that the corresponding deterministic analogue be explicitly solvable, yields results in excellent coincidence with those of the random walk. There is no threshold theorem for rumours as for the general epidemic, since for rumours the effective removal rate equals half the population size.

Daley and Kendall extend their model by introducing a probability α that a spreader becomes a stifler after having met someone who knows the rumour already, and by allowing that a spreader tells the rumour only with probability p when meeting another individual. Another modification supposes that an active spreader of the rumour is removed whenever he encounters k people who have heard it before

(*k*-fold stifling). They also indicate that Feller (1957, p. 55, Exercise 21) gives the distribution of the number of spreaders at the instant where the first meeting of two spreaders occurs, both for their original model and for a model in which people congregate *r* at a time.

ACKNOWLEDGEMENTS

I wish to thank the editor for the invitation to write this review, and I am grateful to Professor J. M. Gani who very kindly consented to read and criticize the manuscript. I should like to thank several authors for allowing me to see their papers in draft prior to publication.

REFERENCES

BAILEY, N. T. J. (1953). The total size of a general stochastic epidemic. *Biometrika*, **40**, 177–185.
—— (1957). *The Mathematical Theory of Epidemics*. London: Griffin; New York: Hafner.
—— (1963). The simple stochastic epidemic: A complete solution in terms of known functions. *Biometrika*, **50**, 235–240.
—— (1964a). *The Elements of Stochastic Processes with Applications to the Natural Sciences*. New York: Wiley.
—— (1964b). Some stochastic models for small epidemics in large populations. *Appl. Statist.*, **13**, 9–19.
—— (1967a). A perturbation approximation to the simple stochastic epidemic in a large population. *Biometrika* (to appear).
—— (1967b). The simulation of stochastic epidemics in two dimensions. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).
—— (1967c). Stochastic birth, death and migration processes for spatially distributed populations. *Biometrika* (to appear).
BAIN, A. D. (1963). The growth of demand for new commodities. *J. R. Statist. Soc.* A, **126**, 285–299.
BARTHOLOMAY, A. F. (1964). The general catalytic queue process. In *Stochastic Models in Medicine and Biology* (ed. J. Gurland), pp. 101–144. University of Wisconsin Press.
BARTLETT, M. S. (1949). Some evolutionary stochastic processes. *J. R. Statist. Soc.* B, **11**, 211–229.
—— (1955). *An Introduction to Stochastic Processes with Special Reference to Methods and Applications*. Cambridge: University Press.
—— (1956). Deterministic and stochastic models for recurrent epidemics. *Proc. 3rd Berkeley Symp. Math. Statist. & Prob.*, **4**, 81–109.
—— (1957). Measles periodicity and community size. *J. R. Statist. Soc.* A, **120**, 48–70.
—— (1960a). *Stochastic Population Models in Ecology and Epidemiology*. New York: Wiley; London: Methuen.
—— (1960b). Some stochastic models in ecology and epidemiology. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (ed. I. Olkin *et al.*), pp. 89–96. Stanford, Calif.: Stanford University Press.
—— (1960c). The critical community size for measles in the United States. *J. R. Statist. Soc.* A, **123**, 37–44.
—— (1961). Monte Carlo studies in ecology and epidemiology. *Proc. 4th Berkeley Symp. Math. Statist. & Prob.*, **4**, 39–55.
—— (1964). The relevance of stochastic models for large-scale epidemiological phenomena. *Appl. Statist.*, **13**, 2–8.
—— (1966). Some notes on epidemiological theory. In *Research Papers in Statistics: Festschrift for J. Neyman* (ed. F. N. David), pp. 25–36. New York: Wiley.
BARTON, D. E. and DAVID, F. N. (1966). The random intersection of two graphs. In *Research Papers in Statistics: Festschrift for J. Neyman* (ed. F. N. David), pp. 445–459. New York: Wiley.
BARTON, D. E., DAVID, F. N. and MERRINGTON, M. (1965). A criterion for testing contagion in time and space. *Ann. Hum. Genet.*, **29**, 97–102.
BARTOSZYŃSKI, R. (1967). Branching processes and the theory of epidemics. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).

BARTOSZYŃSKI, R., ŁOŚ, J. and WYCECH-ŁOŚ, M. (1965). Contributions to the theory of epidemics. In *Bernoulli–Bayes–Laplace Anniversary Volume*, pp. 1–8. Berlin: Springer.

BELLMAN, R. and HARRIS, T. E. (1948). On the theory of age-dependent stochastic branching processes. *Proc. Nat. Acad. Sci.*, **34**, 601–604.

BENAYOUN, R. (1964). Sur un modèle stochastique utilisé dans la théorie mathématique des épidémies. *C.R. Acad. Sci. Paris*, **258**, 5789–5791.

BERNOULLI, D. (1760). Essai d'une nouvelle analyse de la mortalité causée par la petite vérole et des avantages de l'inoculation pour la prévenir. *Mém. Math. Phys. Acad. Roy. Sci., Paris*, 1–45.

BHARUCHA-REID, A. T. (1956). On the stochastic theory of epidemics. *Proc. 3rd Berkeley Symp. Math. Statist. & Prob.*, **4**, 111–119.

—— (1957). *An Introduction to the Stochastic Theory of Epidemics and Some Related Statistical Problems.* Randolph Field, Texas: U.S.A.F. School of Aviation Medicine.

—— (1958). Comparison of populations whose growth can be described by a branching stochastic process—with special reference to a problem in epidemiology. *Sankhyā*, **19**, 1–14.

—— (1960). *Elements of the Theory of Markov Processes and their Applications.* New York: McGraw-Hill.

BLACK, F. L. (1966). Measles endemicity in insular populations: Critical community size and its evolutionary implication. *J. Theor. Biol.*, **11**, 207–211.

BRAMBILLA, F. (1960). Modelli deterministici e stocastici in epidemiologia. *Boll. Centro Ric. Operat. (Serie Metodol.)*, **4**, 3–28.

BÜHLER, W. (1966). A theorem concerning the extinction of epidemics. *Biom. Zeit.*, **8**, 10–14.

CANE, V. R. (1966). On the size of an epidemic and the number of people hearing a rumour. *J. R. Statist. Soc.* B, **28**, 487–490.

DALEY, D. J. (1967). Concerning the spread of news in a population of individuals who never forget. *Bull. Math. Biophys.*, **29**, 373–376.

DALEY, D. J. and KENDALL, D. G. (1964). Epidemics and rumours. *Nature*, **204**, 1118.

—— (1965). Stochastic rumours. *J. Inst. Maths Applics*, **1**, 42–55.

D'ANCONA, U. (1954). *The Struggle for Existence.* Leiden: Brill.

DANIELS, H. E. (1967). The total size of a stochastic epidemic. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).

DIETZ, K. (1966). On the model of Weiss for the spread of epidemics by carriers. *J. Appl. Prob.*, **3**, 375–382.

DIETZ, K. and DOWNTON, F. (1967). Carrier-borne epidemics with immigration. *J. Appl. Prob.* (to appear).

DOWNTON, F. (1967a). A note on the ultimate size of a general stochastic epidemic. *Biometrika*, **54**, 314–316.

—— (1967b). Epidemics with carriers: A note on a paper of Dietz. *J. Appl. Prob.*, **4**, 264–270.

—— (1967c). The size of carrier-borne epidemics. *Biometrika* (to appear).

EDERER, F., MEYERS, M. H. and MANTEL, N. (1964). A statistical problem in space and time: Do leukemia cases come in clusters? *Biometrics*, **20**, 626–638.

ELVEBACK, L., FOX, J. P. and VARMA, A. (1964). An extension of the Reed–Frost epidemic model for the study of competition between viral agents in the presence of interference. *Amer. J. Hyg.*, **80**, 356–364.

ERDÉLYI, A., ed. (1953). *Higher Transcendental Functions*, Volume 1. New York: McGraw-Hill.

FELLER, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheor.*, **5**, 11–40.

—— (1957). *An Introduction to Probability Theory and its Applications*, Volume 1 (2nd ed.). New York: Wiley.

FIX, E. (1967). Space–time contagion. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).

FRISCH, H. L. and HAMMERSLEY, J. M. (1963). Percolation processes and related topics. *J. Soc. Indust. Appl. Math.*, **11**, 894–918.

GAFFEY, W. R. (1954). The probability of within-family contagion. (Thesis, University of California, Berkeley.)

GANI, J. (1965). On a partial differential equation of epidemic theory: I. *Biometrika*, **52**, 617–622.

—— (1966). On the stochastic epidemic with immigration. *Proc. Int. Cong. Mathematicians* (to appear).

—— (1967). On the general stochastic epidemic. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).

GART, J. J. (1965). Some stochastic models relating time and dosage in response curves. *Biometrics*, **21**, 583–599.

GART, J. J. (1967). The mathematical analysis of an epidemic with two kinds of susceptibles (to appear).
GART, J. J. and DE VRIES, J. L. (1966). The mathematical analysis of concurrent epidemics of yaws and chickenpox. *J. Hyg., Camb.*, **64**, 431–439.
GOFFMAN, W. (1965). An epidemic process in an open population. *Nature*, **205**, 831–832.
—— (1966a). A mathematical model for describing the compatibility of infectious diseases. *J. Theor. Biol.*, **11**, 349–361.
—— (1966b). Mathematical approach to the spread of scientific ideas—The history of mast cell research. *Nature*, **212**, 449–452.
GOFFMAN, W. and NEWILL, V. A. (1964). Generalization of epidemic theory. *Nature*, **204**, 225–228.
HAMMERSLEY, J. M. (1966). First-passage percolation. *J. R. Statist. Soc.* B, **28**, 491–496.
HAMMERSLEY, J. M. and WELSH, D. J. A. (1965). First-passage percolation, subadditive processes, stochastic networks and generalized renewal theory. In *Bernoulli–Bayes–Laplace Anniversary Volume*, pp. 61–110. Berlin: Springer.
HARRIS, T. E. (1963). *The Theory of Branching Processes.* Berlin: Springer.
HASKEY, H. W. (1954). A general expression for the mean in a simple stochastic epidemic. *Biometrika*, **41**, 272–275.
HILL, B. M. (1963). The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. *J. Amer. Statist. Ass.*, **58**, 72–84.
HORIUCHI, K. and SUGIYAMA, H. (1957). On the importance of the Monte Carlo approach in the research of epidemiology. *Osaka City Med. J.*, **4**, 59–62.
IGLEHART, D. L. (1964). Multivariate competition processes. *Ann. Math. Statist.*, **35**, 350–361.
IRWIN, J. O. (1963). The place of mathematics in medical and biological statistics. *J. R. Statist. Soc.* A, **126**, 1–41.
KARLIN, S. (1966). *A First Course in Stochastic Processes.* New York: Academic Press.
KEMENY, J. G. and SNELL, J. L. (1960). *Finite Markov Chains.* Princeton: Van Nostrand.
KENDALL, D. G. (1948). On the role of variable generation time in the development of a stochastic birth process. *Biometrika*, **35**, 316–330.
—— (1956). Deterministic and stochastic epidemics in closed populations. *Proc. 3rd Berkeley Symp. Math. Statist. & Prob.*, **4**, 149–165.
—— (1957). La propagation d'une épidémie ou d'un bruit dans une population limitée. *Publ. Inst. Statist. Paris*, **6**, 307–311.
—— (1965). Mathematical models of the spread of infection. In *Mathematics and Computer Science in Biology and Medicine*, pp. 213–225. London: H.M.S.O.
KERMACK, W. O. and MCKENDRICK, A. G. (1927). A contribution to the mathematical theory of epidemics: I. *Proc. Roy. Soc.* A, **115**, 700–721.
KNOX, E. G. (1964). The detection of space–time interactions. *Appl. Statist.*, **13**, 25–29.
—— (1965). Recognition of outbreaks of acute leukaemia and congenital malformations. In *Mathematics and Computer Science in Biology and Medicine*, pp. 227–233. London: H.M.S.O.
LANDAHL, H. D. (1953). On the spread of information with time and distance. *Bull. Math. Biophys.*, **15**, 367–381.
LANDAU, H. G. and RAPOPORT, A. (1953). Contribution to the mathematical theory of contagion and spread of information: I. Spread through a thoroughly mixed population. *Bull. Math. Biophys.*, 173–183.
LOTKA, A. J. (1923). Contributions to the analysis of malaria epidemiology. *Amer. J. Hyg.*, **3** (*Suppl.* 1), 1–121.
MACDONALD, G. (1957). *The Epidemiology and Control of Malaria.* London: Oxford University Press.
MAIA, J. DE O. C. (1952). Some mathematical developments on the epidemic theory formulated by Reed and Frost. *Hum. Biol.*, **24**, 167–200.
MANSFIELD, E. and HENSLEY, C. (1960). The logistic process; tables of the stochastic epidemic curve and application. *J. R. Statist. Soc.* B, **22**, 332–337.
MARCHAND, H. (1956). Essai d'étude mathématique d'une forme d'épidémie. *Ann. Univ. Lyon* A(3), **19**, 13–46.
MCKENDRICK, A. G. (1926). Applications of mathematics to medical problems. *Proc. Edinb. Math. Soc.*, **44**, 98–130.
MEYNELL, G. G. and MEYNELL, E. W. (1958). The growth of micro-organisms *in vivo* with particular reference to the relation between dose and latent period. *J. Hyg., Camb.*, **56**, 323–346.
MORGAN, R. W. (1964). A note on Dr Bailey's paper. *Appl. Statist.*, **13**, 20–24.
—— (1965). The estimation of parameters from the spread of a disease by considering households of two. *Biometrika*, **52**, 271–274.

MORGAN, R. W. and WELSH, D. J. A. (1965). A two-dimensional Poisson growth process. *J. R. Statist. Soc.* B, **27**, 497–504.
MUENCH, H. (1959). *Catalytic Models in Epidemiology*. Boston, Mass.: Harvard University Press.
NEYMAN, J. (1965). Certain chance mechanisms involving discrete distributions. *Sankhyā* A, **27**, 249–258.
NEYMAN, J. and SCOTT, E. (1964). A stochastic model of epidemics. In *Stochastic Models in Medicine and Biology* (ed. J. Gurland), pp. 45–85. University of Wisconsin Press.
OHLSEN, S. (1964). On estimating epidemic parameters from household data. *Biometrika*, **51**, 511–512.
PETTIGREW, H. M. and WEISS, G. H. (1967). Epidemics with carriers: The large population approximation. *J. Appl. Prob.*, **4**, 257–263.
PURI, P. S. (1966). On the homogeneous birth-and-death process and its integral. *Biometrika*, **53**, 61–71.
—— (1967). A class of stochastic models of response after infection in the absence of defense mechanism. *Proc. 5th Berkeley Symp. Math. Statist. & Prob.* (to appear).
PYKE, R. (1965). Spacings. *J. R. Statist. Soc.* B, **27**, 395–449.
RAPOPORT, A. (1953a). Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *Bull. Math. Biophys.*, **15**, 523–533.
—— (1953b). Spread of information through a population with socio-structural bias: II. Various models with partial transitivity. *Bull. Math. Biophys.*, **15**, 535–546.
RAPOPORT, A. and REBHUN, L. J. (1952). On the mathematical theory of rumor spread. *Bull. Math. Biophys.*, **14**, 375–383.
REUTER, G. E. H. (1957). Denumerable Markov processes and the associated contraction semigroups on *l*. *Acta Math.*, **97**, 1–46.
—— (1961). Competition processes. *Proc. 4th Berkeley Symp. Math. Statist. & Prob.*, **2**, 421–430.
RIDLER-ROWE, C. J. (1967). On a stochastic model of an epidemic. *J. Appl. Prob.*, **4**, 19–33.
SAATY, T. L. (1961a). Some stochastic processes with absorbing barriers. *J. R. Statist. Soc.* B, **23**, 319–334.
—— (1961b). *Elements of Queueing Theory*. New York and London: McGraw-Hill.
SAKINO, S. (1959). Epidemic model under the influence of damped oscillation. *Proc. Inst. Statist. Math.*, **6**, 117–124. (In Japanese.)
—— (1962a). Historical development of epidemic models: 1. *Proc. Inst. Statist. Math.*, **9**, 127–136. (In Japanese.)
—— (1962b). On the analysis of epidemic model: II. Theory and application. *Ann. Inst. Statist. Math.*, **13**, 147–163.
—— (1962c). On the age distribution in epidemic. *Proc. Inst. Statist. Math.*, **10**, 33–39. (In Japanese.)
—— (1963). On the age distribution of epidemic model and the prediction of poliomyelitis. *Proc. Inst. Statist. Math.*, **11**, 25–30. (In Japanese.)
SAKINO, S. and HAYASHI, C. (1959). On the analysis of epidemic model: I. Theoretical approach. *Ann. Inst. Statist. Math.*, **10**, 261–275.
SARTWELL, P. E. (1950). The distribution of incubation periods of infectious disease. *Amer. J. Hyg.*, **51**, 310–318.
SCHWÖBEL, W., GEIDEL, H. and LORENZ, R. J. (1966). Ein Modell der Plaquebildung. *Zeit. Naturf.*, **21**(*b*), 953–960.
SEIDEN, E. (1957). On a mathematical model for a problem in epidemiology. *Bull. Amer. Math. Soc.*, **63**, 142–143.
SERFLING, R. E. (1952). Historical review of epidemic theory. *Hum. Biol.*, **24**, 145–166.
SEVERO, N. C. (1967a). Two theorems on solutions of differential–difference equations and applications to epidemic theory. *J. Appl. Prob.*, **4**, 271–280.
—— (1967b). The probabilities of a carrier epidemic model with arbitrary initial distribution (to appear).
SISKIND, V. (1965). A solution of the general stochastic epidemic. *Biometrika*, **52**, 613–616.
SPICER, C. C. and LIPTON, S. (1958). Numerical studies on some contagious distributions. *J. Hyg., Camb.*, **56**, 516–522.
STÖRMER, H. (1964). Das Anwachsen des Verbrauchs von Wirtschaftsgütern als stochastischer Prozess. *Zeit. angew. Math. Mech.*, **44**, 72–73.
SUGIYAMA, H. (1961). Some statistical methodologies for epidemiological research of medical sciences. *Bull. Int. Statist. Inst.*, **38** (3), 137–151.
TAYLOR, I. and KNOWELDEN, J. (1964). *Principles of Epidemiology* (2nd ed). London: Churchill.

TAYLOR, W. F. (1956). Problems in contagion. *Proc. 3rd Berkeley Symp. Math. Statist. & Prob.*, **4**, 167–179.
—— (1958). Some Monte Carlo methods applied to an epidemic of acute respiratory disease. *Hum. Biol.*, **30**, 185–200.
WEISS, G. H. (1965). A model for the spread of epidemics by carriers. *Biometrics*, **21**, 481–490.
WHITTLE, P. (1955). The outcome of a stochastic epidemic—A note on Bailey's paper. *Biometrika*, **42**, 116–122.
WILLIAMS, T. (1965a). The simple stochastic epidemic curve for large populations of susceptibles. *Biometrika*, **52**, 571–579.
—— (1965b). The distribution of response times in a birth–death process. *Biometrika*, **52**, 581–585.
—— (1965c). The basic birth–death model for microbial infections. *J. R. Statist. Soc.* B, **27**, 338–360.
WILLIAMS, T. and MEYNELL, G. G. (1967). Time-dependence and count-dependence in microbial infection. *Nature*, **214**, 473–475.
YANG, G. L. (1966). Contagion in stochastic models for epidemics. (Thesis, University of California, Berkeley.)