Adv. Appl. Prob. **25**, 303–313 (1993)
Printed in N. Ireland

DYNAMICS OF AN EPIDEMIC IN A CLOSED POPULATION

ÅKE SVENSSON, **Stockholm University*

Abstract

A simple model for the intensity of infection during an epidemic in a closed population is studied. It is shown that the size of an epidemic (i.e. the number of persons infected) and the cumulative force of an epidemic (i.e. the amount of infectiousness that has to be avoided by a person that will stay uninfected during the entire epidemic) satisfy an equation of balance. Under general conditions, small deviances from this balance are, in large populations, asymptotically mixed normally distributed. For some special epidemic models the size of an asymptotically large epidemic is asymptotically normally distributed.

EPIDEMIC MODELS; COUNTING PROCESSES; MARTINGALE LIMIT THEOREMS; SIZE DISTRIBUTION

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60K30
SECONDARY 60G44, 60F05, 92C99

1. Introduction

We consider a population that at time $t = 0$ consists of W healthy individuals. At this time an infectious disease is introduced into the population, and as time goes on individuals are infected. They then spread the disease according to some biological mechanism that depends on the specific character of the particular infection: in other words an epidemic starts. We will study a class of stochastic models for the development of such epidemics. The models are studied within the framework of counting processes and martingales.

There is a substantial mathematical theory for the spread of infectious diseases. The mathematical methods used include non-stochastic models based on differential–difference equations, simulation models and stochastic models built on Markov processes (cf. Bailey (1975), Anderson (1982) and Lefèvre (1990)).

We define some fundamental concepts and derive a general 'equation of balance' for the relation between the size and the force of an epidemic. We also study the asymptotic behaviour of the size of some special epidemics. Some of the results obtained are already known, and proved by other methods than the one used here. One of the aims of this paper is to illustrate a technique that may be used to derive interesting results.

Received 22 May 1990; revision received 8 May 1992.
* Postal address; Department of Statistics, Stockholm University, 5106 91 Stockholm, Sweden.

2. A counting process model for an epidemic

Let $N(t)$ be the number of individuals in the population that have been infected up till time t. The process N is assumed to be adapted to a filtration $\mathscr{A} = \{\mathscr{A}_t\}_{t=0}^{\infty}$, and to have an intensity function λ that is predictable relative to $\mathscr{A}$. The filtration has to be defined for every epidemic model. Typically $\mathscr{A}_t$ is generated by the history of the process N and by random variables that describe the infectiousness of the infected individuals up until time t.

$N(t)$ is a counting process, and we will assume that its intensity function has the form $\lambda(t) = (W - N(t-))\xi(t)$. The intensity is the product of two factors. The first factor $W - N(t-)$ is the number of susceptible individuals which are at risk of acquiring the infection immediately before time t. The function $\xi(t)$ will be called the force of the epidemic at $t-$, and measures the amount of infectivity that a susceptible individual will meet at time t. It will depend on the previous history of the epidemic, i.e. how much infectious matter is introduced into the population from outside and how much infectivity is generated by individuals inside the population. Thus, it is a random function that is predictable relative to $\mathscr{A}$. ξ will decide the future of the epidemic. An important implicit assumption is that the spread of the infection is homogeneous in the population, i.e. the progress of the epidemic is not influenced by which particular persons are infected.

In the following we will use the fact that $N(t) - \int_0^t \lambda(s)\, ds$ is a martingale relative to $\mathscr{A}$. Other (local) martingales can be formed as integrals $\int_0^1 r(s)(dN(s) - \lambda(s)\, ds)$ where r is a predictable process. If $\tau_1, \tau_2, \cdots$ are the successive times when individuals become infected and θ_i, $i = 1, 2, \cdots$ are $\mathscr{A}_{\tau_i}$-measurable random variables such that $\boldsymbol{E}(\theta_i \mid \mathscr{A}_{\tau_i -}) = 0$ then $\sum_{i=1}^{N(t)} \theta_i$ is a martingale.

First some examples of epidemic models. In all cases it is assumed that the epidemic is influenced by an external infectious force, denoted by $B(t)$.

Example 1. An individual that is infected will start spreading the disease after a latent (constant) time Δ_1 and will then be infectious during a time interval of (constant) length $\Delta_2 - \Delta_1$. During this time his 'infectiousness' will be constant equal to β. At time $t-$ there will be $N(t - \Delta_1) - N(t - \Delta_2)$ infectious persons in the population. If the infectiousness is spread homogeneously over the population then

$$\xi(t) = B(t)/W + \beta(N(t - \Delta_1) - N(t - \Delta_2))/W.$$

The total amount of infectivity spread by one infected person is in this model $\alpha = \beta(\Delta_2 - \Delta_1)$.

The Reed–Frost model is a discrete-time version of this example. The time scale is divided into intervals of equal length. For simplicity this length can be assumed to equal one unit of time. The epidemic starts at time $t = 0$ with m infected persons entering the population from outside. They will be infectious for one unit of time. An infected person starts to be infectious at the start of the time interval following his infection and will be infectious during one unit of time. With $B(t) = \beta m$ when

$t \in [0, 1[$ and 0 if $t \geqq 1$ this epidemic will be defined by $\xi(t) = B(t)/W + \beta(N([t]) - N([t-1]))/W$, ($[t]$ is the integer part of t). The Reed–Frost epidemic has been thoroughly studied. Many of the results obtained in this paper are already known to hold for this epidemic process (cf. von Bahr and Martin-Löf (1980)).

Example 2. After being infected an individual is infectious according to some function $f(u)$ where u is the time since infection. This gives

$$\xi(t) = B(t)/W + \int_0^{t-} f(t-s)\, dN(s)/W.$$

This is a slight generalization of Example 1. The total amount of infectivity spread by one infected individual is $\alpha = \int_0^\infty f(u)\, du$.

Example 3: IIDI epidemics. Several models have been suggested where it is assumed that the amount of infectivity admitted from one infected individual is not constant but a random variable. In the stochastic version of the so-called Kermack–McKendrick model an infected individual is assumed to spread the infection with a constant intensity during a random time interval. The lengths of these time intervals are assumed to be independently identically distributed random variables. In the simplest case the distribution is assumed to be exponential (cf. Frauenthal (1980)). This model is sometimes referred to as the general (or standard) epidemic model (cf. Watson (1981)). Kurtz (1981) assumes a general form of the distribution. We will assume that the total amount of infectivity spread by the ith infected person is $\alpha_i = \int_0^\infty f_i(u)\, du$. The α_i's are assumed to be independent identically distributed random variables with mean $\alpha = \boldsymbol{E}(\alpha_i)$ and finite variance σ^2. We will call epidemic models of this kind IIDI (independent identically distributed infectivities). The randomized Reed–Frost models studied by von Bahr and Martin-Löf (1980) can be viewed as a discrete-time version.

Example 4: Daley–Kendall and Maki–Thompson rumour models. Daley and Kendall (1965) suggested an epidemic model for the spread of rumours. They assume that at time $t = 0$ one person starts spreading a rumour in a population of size W. An individual that has heard of the rumour tells it until he encounters a person who already knows of the rumour. The behaviour of epidemic models of this kind have been studied by Watson (1988) and Pittel (1990). Maki and Thompson (1973) suggested a modification of this model. They assume that if two active spreaders meet, only one of them stops spreading the rumour (cf. Watson (1988)).

3. The size and cumulative force of an epidemic

We will derive asymptotic properties of the epidemic which are valid in large populations, i.e. when $W \to \infty$. To indicate the dependence on the size of the population we use the index W.

The process $N_W(t)$ is increasing and bounded by W. There thus exists a limit

$$S_W = \lim_{t\to\infty} N_W(t),$$

which is the final size of the epidemic, i.e. the total number of persons infected before the epidemic dies out.

Another entity of importance is the cumulative force of the epidemic, i.e.

$$C_W = \int_0^\infty \xi_W(t)\,dt.$$

C_W can be interpreted as the total amount of infectiousness encountered by a single individual who has stayed non-infected during the entire epidemic. In the case where all individuals are eventually infected, we also need to consider the effective cumulative force of the epidemic. This number is defined as

$$C_W^* = \int_0^\infty \xi_W(t) I(N_W(t) < W)\,dt.$$

Of course, S_W and C_W are intimately related. In Examples 1 and 2 the relation

$$C_W = B/W + \alpha S_W/W \tag{3.1}$$

will hold (provided $S_W < W$). Here $B = \int_0^\infty B(t)\,dt$. An epidemic model where relation (3.1) holds is called a linear epidemic.

For an IIDI epidemic (cf. Example 3) the relation $C_W = B/W + \sum_1^{S_W} \alpha_i/W$ holds. Unless the distribution for the α_i's is degenerate, this model will not be linear. From Kolmogorov's inequality it follows that $\sum_{i=1}^{S_W} (\alpha_i - \alpha)/W \xrightarrow{\text{P}} 0$ as $W \to \infty$. Thus an IIDI epidemic satisfies the relation

$$C_W - \alpha S_W/W \xrightarrow{\text{P}} 0 \tag{3.2}$$

as $W \to \infty$. An epidemic model with this property will be called asymptotically proportional. The Maki–Thompson rumour model is also an asymptotically proportional epidemic model with $\alpha = 2$ (see Section 5.2).

4. Some limit theorems

Lemma 4.1. Let $M_W(s)$ be a sequence of zero-mean martingales, such that $\boldsymbol{E}([M_W](s)) \leqq K_W$, where $K_W \to 0$, for all $s \in [0, \infty[$, then $M_W(\infty) \xrightarrow{\text{P}} 0$ as $W \to \infty$.

Proof. If a martingale M has uniformly bounded second moments, i.e. $\boldsymbol{E}(M^2(s)) = \boldsymbol{E}(\langle M\rangle(s)) = \boldsymbol{E}([M](s)) \leqq K < \infty$ for all $s \in [0, \infty[$, then $\lim_{s\to\infty} M(s) = M(\infty)$ exists almost surely and $\boldsymbol{E}(M^2(\infty)) \leqq K$. (cf. Jacod and Shiryaev (1987), Theorem 1.42).

Since $\boldsymbol{E}(M_W^2(\infty)) \leqq K_W \to 0$, the lemma follows from Chebyshev's inequality.

Lemma 4.2. $\sum_{n=0}^{S_W-1} 1/(W-n) - C_W^*$ has uniformly bounded variances.

Proof. The martingale

$$M_W(t) = \int_0^t dN_W(s)/(W - N(s-)) - \int_0^t \xi_W(s) I(N_W(s-) < W)\, ds$$

is square integrable with optional quadratic variation

$$[M_W](t) = \int_0^t dN_W(s)/(W - N_W(s-))^2 \leqq \sum_{n=0}^{W-1} 1/(W-n)^2 \leqq \sum_1^\infty 1/n^2 < \infty.$$

Now

$$M_W(\infty) = \sum_{n=0}^{S_W - 1} 1/(W-n) - C_W^*.$$

Since $E(M_W^2(\infty)) = E([M_W](\infty)) \leqq \sum_1^\infty 1/n^2 < \infty$ the lemma follows.

Theorem 4.3. If $\ln(W) - C_W \xrightarrow{P} \infty$ as $W \to \infty$, then $W - S_W \xrightarrow{P} \infty$ and $\ln(1 - S_W/W) + C_W \xrightarrow{P} 0$ as $W \to \infty$.

Proof. First observe that

$$\int_{W-S_W}^W dx/x \geqq \sum_{n=0}^{S_W-1} 1/(W-n) \geqq \int_{W-S_W}^W dx/x - \int_{W-S_W}^W dx/x^2.$$

From Lemma 4.2 and Chebyshev's inequality it follows that $\sum_{n=0}^{S_W-1} 1/(W-n) - C_W^* \geqq \ln(W) - C_W - \ln(W - S_W) - S_W/(W(W - S_W))$ is bounded from above in probability. If $\ln(W) - C_W \xrightarrow{P} \infty$, this can only happen if $W - S_W \xrightarrow{P} \infty$. Furthermore $C_W - C_W^* \xrightarrow{P} 0$, since these two numbers differ only when $W = S_W$.

If $W - S_W \xrightarrow{P} \infty$

$$[M_W](t) = \int_0^t dN_W(s)/(W - N_W(s-))^2$$

$$\leqq \sum_{n=0}^{S_W-1} 1/(W-n)^2 \leqq I(S_W = W) + 1/(W - S_W + 1) \xrightarrow{P} 0$$

as $W \to \infty$.

By Lemma 4.1 it follows that $M_W(\infty) \xrightarrow{P} 0$. Another consequence of the inequality above is that

$$\sqrt{W}\left(\sum_{n=0}^{S_W-1} 1/(W-n) + \ln(1 - S_W/W)\right) \xrightarrow{P} 0 \tag{4.1}$$

and thus $\ln(1 - S_W/W) + C_W \xrightarrow{P} 0$ as $W \to \infty$.

Theorem 4.4. If $S_W/W \xrightarrow{\mathscr{L}} P$ where P is a random variable with support on $[0, K[$ with $K < 1$, and if there exists a sequence, η_W, of stopping times such that

(i) $N(\eta_W)/W \to 0$ and

(ii) $E(S_W/W \mid \mathscr{A}_{\eta_W}) - S_W/W \xrightarrow{P} 0$

then $\sqrt{W}(\ln(1-S_W/W)+C_W)\xrightarrow{\mathscr{L}} Z\sqrt{P/(1-P)}$ as $W\to\infty$, where Z is a N(0, 1)-distributed random variable that is independent of P.

Proof. We start by defining the stopping time

$$\tau_W=\inf\{t;\, N_W(t)/W\geqq(1+K)/2\}.$$

Since the asymptotic distribution of S_W/W is concentrated on $[0, K[$ it follows that τ_W equals ∞ with a probability that tends to 1 as $W\to\infty$.

The martingales

$$M_W^{(r)}(t)=W^{r/2}\int_0^t(dN_W(s)-\lambda_W(s)\,ds)/(W-N_W(s-))^r$$

have optional quadratic variations

$$[M_W^{(r)}](t)=W^r\int_0^t dN_W(s)/(W-N_W(s-))^{2r}\leqq W^{-r}N_W(t)(1-N_W(t)/W)^{-2r}.$$

Applying Lemma 4.1 to the stopped martingales $M_W^{(r)}(t\wedge\tau_W)$ it follows that $M_W^{(r)}(\tau_W)\xrightarrow{\mathrm{P}}0$ as $W\to\infty$ if $r>1$.

Define the cumulative function

$$K_W(\mu,t)=\int_0^t(\exp\{i\mu/(W-N_W(s-))\}-1-i\mu/(W-N_W(s-)))\lambda_W(s)\,ds, \tag{4.2}$$

then $Z_W(\mu,t)=\exp\{i\mu M_W(t)-K_W(\mu,t)\}$ is the Doléans–Dade exponential of the martingale $\int_0^t(\exp\{i\mu/(W-N_W(s-))\}-1)(dN_W(s)-\lambda_W(s)\,ds)$ (cf. Jacod and Shiryaev (1987)). This means that

$$\begin{aligned}\tilde{Z}_W(\mu,t)&=Z_W(\mu,t)-1\\ &=\int_0^t Z_W(\mu,s)(\exp\{i\mu/(W-N_W(s-))\}-1)(dN_W(s)-\lambda_W(s)\,ds).\end{aligned}\tag{4.3}$$

This follows from general theory, but is also easily checked in this particular case. Thus $\tilde{Z}_W(\mu,t)$ is, for any μ, a local zero-mean martingale.

From the elementary inequality $|\exp\{ix\}-1-ix+x^2/2|\leqq|x|^3/6$ it follows that

$$\begin{aligned}&\left|K_W(\sqrt{W}\mu,t)+W(\mu^2/2)\int_0^t 1/(W-N_W(s-))^2\,dN_W(s)\right|\\ &\leqq\left|K_W(\sqrt{W}\mu,t)+W(\mu^2/2)\int_0^t 1/(W-N_W(s-))^2\lambda_W(s)\,ds\right|+(\mu^2/2)\,|M_W^2(t)|\\ &\leqq(\mu^2/2)\,|M_W^{(2)}(t)|+|\mu^3/6|\,|M_W^{(3)}(t)|+W^{\frac{3}{2}}\,|\mu^3/6|\int_0^t 1/(W-N_W(s-))^3\,dN_W(s).\end{aligned}$$

If

$$\sigma_W=\inf\{t;\,|M_W^{(2)}(t)|>1,\,|M_W^{(3)}(t)|>1\}\wedge\tau_W$$

then it follows from the inequality

$$P\Big(\sup_{t\leqq\tau_W}|M_W^{(r)}(t)|\geqq 1\Big)\leqq E([M_W^{(r)}](\tau_W))$$

that $\sigma_W=\tau_W$ with a probability that tends to 1 as $W\to\infty$, i.e. σ_W equals ∞ asymptotically.

We can conclude that $K_W(\sqrt{W}\,\mu,\,\sigma_W)$ is uniformly bounded in W and asymptotically equivalent to $-W(\mu^2/2)\int_0^{\sigma_W}dN_W(s)/(W-N_W(s-))^2$. These two random variables are also asymptotically equivalent to $-W(\mu^2/2)\int_0^{\infty}dN_W(s)/(W-N_W(s-))^2$ and $-(\mu^2/2)S_W/(W-S_W)$. If $P_W=E(S_W/W\mid\mathscr{A}_{\eta_W})$ then it is also asymptotically equivalent to $-(\mu^2/2)P_W/(1-P_W)$ due to condition (ii).

By dominated convergence it follows that

$$E(\exp\{K_W(\sqrt{W}\,\mu,\,\sigma_W)\})\to E(\exp\{-(\mu^2/2)P/(1-P)\}).$$

Observe that this is the characteristic function of the random variable $Z\sqrt{P/(1-P)}$, where Z is N(0, 1)-distributed and independent of P.

Now

$$\begin{aligned}
&E(\exp\{i\sqrt{W}\,\mu M_W(\sigma_W)\})-E(\exp\{K_W(\sqrt{W}\,\mu,\,\sigma_W)\})\\
&=E(\exp\{K_W(\sqrt{W}\,\mu,\,\sigma_W)\}\tilde{Z}_W(\sqrt{W}\,\theta,\,\sigma_W))\\
&=E((\exp\{K_W(\sqrt{W}\,\mu,\,\sigma_W)\}\\
&\quad-\exp\{-(\mu^2/2)P_W/(1-P_W)\})\tilde{Z}_W(\sqrt{W}\,\mu,\,\sigma_W))\\
&\quad+E(E(\exp\{-(\mu^2/2)P_W/(1-P_W)\}\tilde{Z}_W(\sqrt{W}\,\mu,\,t)\mid\mathscr{A}_{\eta_W})).
\end{aligned}\tag{4.4}$$

The first term of (4.4) tends to 0 as $W\to\infty$ due to the asymptotic equivalence proved above. The second term equals $E(\exp\{-(\mu^2/2)P_W/(1-P_W)\}\tilde{Z}_W(\sqrt{W}\,\mu,\,\eta_W))$ since P_W is $\mathscr{A}_{\eta_W}$-measurable and $\tilde{Z}_W(\sqrt{W}\,\mu,\,\cdot)$ is a martingale. Due to (4.3) and since $K_W(\sqrt{W},\,\eta_W)$ is bounded

$$[\tilde{Z}_W(\sqrt{W}\,\mu;\,\cdot)](\eta_W)\leqq\text{const.}\;\mu^2N_W(\eta_W)/(W-N_W(\eta_W)).$$

This random variable tends to 0 as $W\to\infty$ due to condition (i) and also the second term of (4.4) is asymptotically small. Thus $\sqrt{W}\,M_W(\infty)\xrightarrow{\mathscr{L}}Z\sqrt{P/(1-P)}$. This together with (4.1) proves the theorem.

This theorem implies that $\sqrt{W}(\ln(1-S_W/W)+C_W)$ is asymptotically mixed normally distributed.

5. Limit theorems for special epidemic models

Theorem 5.1. In an asymptotically proportional epidemic model S_W/W is asymptotically concentrated on the (at most two) roots of the equation

$$\alpha\pi+\ln(1-\pi)=0.\tag{5.1}$$

Proof. In an asymptotically proportional epidemic model, $C_W - \alpha S_W/W \xrightarrow{P} 0$ as $W \to \infty$. Since $S_W \leqq W$ it follows that C_W is asymptotically bounded in probability. Thus, we can apply Theorem 4.3. It follows that $\ln(1 - S_W/W) + \alpha S_W/W \xrightarrow{P} 0$. The equation $\ln(1-\pi) + \alpha\pi = 0$ has only the root $\pi = 0$ if $\alpha \leqq 1$. If $\alpha > 1$ it also has one positive root. The distribution of S_W/W has to be asymptotically concentrated on these roots.

Lemma 5.2. In any asymptotically proportional epidemic model, $I(S_W \leqq \delta W)\sqrt{W}\, M_W(\infty) \xrightarrow{P} 0$ and in a linear epidemic, with $\alpha \neq 1$, $I(S_W \leqq \delta W) S_W/\sqrt{W} \xrightarrow{P} 0$ as $W \to \infty$, for any δ such that $0 < \delta < \pi$, where π is the largest root of (5.1).

Proof. Define the stopping times $\eta_W = \inf\{t; N_W(t) > \varepsilon W\}$. Since

$$W[M_W](\eta_W) = W\int_0^{\eta_W} dN_W(s)/(W - N_W(s-))^2 \leqq \varepsilon/(1-\varepsilon)$$

it follows from Chebyshev's inequality that

$$\boldsymbol{P}(\sqrt{W}\, M_W(\eta_W) \geqq \varepsilon^{\frac{1}{4}}) \to 0$$

as $W \to \infty$ for any $\varepsilon > 0$. For small epidemics $\sqrt{W}\, M_W(\eta_W) = \sqrt{W}\, M_W(\infty)$. If δ is any number strictly less than the positive root of (5.1) $I(S_W/W \leqq \delta) \to 0$ if the epidemic is asymptotically large. Thus the first part of the lemma is proved.

For a linear epidemic $\sqrt{W}\,(M_W(\infty) + \ln(1 - S_W/W) + \alpha S_W/W) \xrightarrow{P} 0$. If S_W/W is small $\ln(1 - S_W/W) + \alpha S_W/W \approx (\alpha - 1)S_W/W$. This implies the second part of the lemma.

Theorem 5.3. In a linear epidemic with $\alpha > 1$ the conditional distribution of $\sqrt{W}\,(S_W/W - \pi)$ given that the epidemic is large is asymptotically $\mathrm{N}(0, \pi(1-\pi)/(1 - \alpha + \alpha\pi)^2)$, where π is the positive root of (5.1).

Proof. The sequence of stopping times $\eta_W = \inf\{t; N_W(t) > \sqrt{W}\}$ trivially satisfies conditions (i) of Theorem 4.4. It follows from Lemma 5.2 that it also satisfies condition (ii).

From a Taylor expansion, around π, it is seen that

$$\sqrt{W}\,(\ln(1 - S_W/W) + \alpha S_W/W) \approx \sqrt{W}\,(S_W/W - \pi)(-1/(1-\pi) + \alpha) \tag{5.2}$$

when S_W/W is close to π.

From Theorem 4.4 and by identifying the parts of the distribution it is seen that the conditional distribution of (5.2) given that the epidemic is large has to be $\mathrm{N}(0, \pi/(1-\pi))$. The theorem then follows from simple calculations.

5.1. *IIDI epidemics*. To describe these models we will use a filtration where $\mathscr{A}_t$ is generated by the process $N_W(s)$ for $s \leqq t$ and $\alpha_1, \ldots, \alpha_{N_W(t)}$.

Theorem 5.4. In an IIDI epidemic model, with external force B, the asymptotic probability of a small epidemic is $\exp(-B\delta)$ where δ is the smallest positive root of

$$\boldsymbol{E}(\exp(-\alpha_i\delta)) = 1 - \delta. \tag{5.3}$$

Proof. The probability for a small epidemic, $\zeta(B, W)$, depends on the initial force B and the size of the population. This probability will, for fixed B, be a decreasing function of W. There will thus exist a limit, $\zeta(B)$ which is the asymptotic probability for a small epidemic.

For a small epidemic and any $\varepsilon > 0$

$$\left|\sum_{i=1}^{S_W} (\alpha_i - \alpha)/\sqrt{W}\right| \leqq \sup_{n<\varepsilon W} \left|\sum_{i=1}^{n} (\alpha_i - \alpha)/\sqrt{W}\right|$$

with a probability that tends to 1 as $W \to \infty$. By Kolmogorov's inequality the right-hand side of this expression can be made arbitrarily small by choosing ε sufficiently small. Thus

$$I(S_W < \delta W) \sum_{i=1}^{S_W} (\alpha_i - \alpha)/\sqrt{W} \xrightarrow{\text{P}} 0$$

as $W \to \infty$. Since

$$\sqrt{W}\, M_W(\infty) \approx \sqrt{W}\left(-\ln(1 - S_W/W) - \alpha S_W/W - \sum_{i=1}^{S_W} (\alpha_i - \alpha)/W\right)$$

it follows from Lemma 5.2 that $S_W/\sqrt{W}$ tends to 0 as $W \to \infty$ if the epidemic is small.

Let us assume that we can divide the external force into two parts such that $B = B_1 + B_2$. In order for the entire epidemic to be small, both these initial forces will have to generate small epidemics. Let $S_W^{(1)}$ and $S_W^{(2)}$ be the sizes of the two epidemics. The probability that the epidemics overlap, i.e. they reach the same person, is less than $S_W^{(1)} S_W^{(2)}/W$, which will tend to 0 as $W \to \infty$. Thus, the asymptotic equality $\zeta(B_1 + B_2) = \zeta(B_1)\zeta(B_2)$ will hold. Consequently

$$\zeta(B) = \exp\{-B\delta\} \tag{5.4}$$

for some δ. Until the first individual is infected the intensity of the counting process N equals $B(t)$. The probability that no individual in the population will be infected is $\exp\{-B\}$. The amount of infectiousness, b, spent until the first person in the population is infected will be (censored) exponentially distributed. When the first person in the population is infected the (random) infectious force α_1 is added. This will give the equation

$$\zeta(B) = \exp\{-B\} + \int_0^\infty \int_0^B \zeta(B - b + \alpha_1) \exp\{-b\}\, db\, dF(\alpha_1),$$

where F is the distribution function of the α_i's. Together with (5.4) this equation implies Equation (5.3).

Theorem 5.5. For an IIDI epidemic, in which the total amount of infectivity spread by an infected person is a random variable with mean $\alpha > 1$ and variance σ^2, the conditional distribution of $\sqrt{W}(S_W/W - \pi)$ given that the epidemic is large is asymptotically $N(0, \pi(1-\pi)(1+(1-\pi)\sigma^2)/(1-\alpha+\alpha\pi)^2)$, where π is the positive root of (5.1).

Proof. Let $\varphi(\mu) = \boldsymbol{E}(\exp\{i\mu(\alpha_i - \alpha)\})$ be the characteristic function of the random variables $(\alpha_i - \alpha)$. The function

$$X_W(\mu, t) = \exp\left\{i\mu \sum_{i=1}^{N_W(t)} (\alpha_i - \alpha)/W - N_W(t) \ln(\varphi(\mu/W)) + i\mu M_W(t) - K_W(\mu, t)\right\}$$

(with K_W defined by (4.2)) is the Doléans–Dade exponential of the martingale

$$\int_0^t (\exp\{i\mu/(W - N_W(s-))\} - 1)(dN_W(s) - \lambda_W(s)\, ds)$$

$$+ \sum_{n=1}^{N_W(t)} \exp\{i\mu/(W-n)\}(\exp\{i\mu(\alpha_n - \alpha)/W - \ln(\varphi(\mu/W))\} - 1).$$

Thus $X_W(\mu, t) - 1$ is a zero-mean martingale. As $W \to \infty$, $N_W(\infty) \ln(\varphi(\mu/\sqrt{W}))$ is asymptotically equivalent to $-(\mu^2/2)\sigma^2 S_W/W$.

Let $\eta_W = \inf\{t; N_W(t) > \sqrt{W}\}$ then $X_W(\sqrt{W}\mu, \eta_W) \to 1$ as $W \to \infty$. Using the same arguments as in the proof of Theorem 4.4 we can prove that

$$\boldsymbol{E}(\exp\{S_W \ln(\varphi(\mu/\sqrt{W})) - K_W(\sqrt{W}\mu, \infty)\}(X_W(\sqrt{W}\mu, \infty) - 1)) \to 0$$

as $W \to \infty$. This implies that

$$\left(\sum_{i=1}^{S_W} (\alpha_i - \alpha)/\sqrt{W} + \sqrt{W} M_W(\infty)\right) \to Z\sqrt{P\sigma^2 + P/(1-P)} \tag{5.5}$$

in law as $W \to \infty$. Here P is a random variable concentrated at the two solutions of (5.1) and Z is a $N(0, 1)$-distributed random variable independent of P. Now $\sum_{i=1}^{S_W} (\alpha_i - \alpha)/\sqrt{W} + \sqrt{W} M_W(\infty)$ is asymptotically equivalent to $-(\ln(1 - S_W/W) + \alpha S_W/W)$. Finally, a Taylor expansion around π and the convergence (5.5) yields the theorem.

Results of the kind proved in Theorems 5.4 and 5.5 are given by von Bahr and Martin-Löf (1980). The asymptotic distribution of the size of a Reed–Frost process is given by Theorem 5.5 with $\sigma^2 = 0$.

5.2. *The Maki–Thompson model.* Maki and Thompson's model for rumour spread is described in Example 4 of Section 2. Let $N_W(t)$ be the number of individuals who have heard the rumour up till time t. We assume that an active spreader in the mean tells the rumour to β individuals per time unit.

The counting process N_W has the intensity

$$\lambda_W(t) = \beta(W - N_W(t-))(N_W(t-) - D_W(t-) + 1),$$

where $D_W(t)$ is the number of individuals who know the rumour but have stopped spreading it at time t. D_W is also a counting process and has the intensity

$$\mu_W(t) = \beta(N_W(t) + 1)(N_W(t-) - D_W(t-) + 1).$$

The cumulative force of the epidemic is

$$C_W = \beta \int_0^\infty (N_W(s) - D_W(s) + 1)\, ds.$$

We can observe that

$$\tilde{M}_W(t) = (N_W(t) + D_W(t))/(W + 1) - \beta \int_0^t (N_W(s) - D_W(s) + 1)\, ds$$

is a zero-mean martingale. It is easy to verify that $\tilde{M}_W$ satisfies the assumptions in Lemma 4.1. Thus $\tilde{M}_W(\infty) \xrightarrow{P} 0$ as $W \to \infty$. Since $N_W(\infty) = D_W(\infty) = S_W$, it follows that $C_W - 2S_W/W \xrightarrow{P} 0$. Thus, we have proved the following lemma.

Lemma 5.6. The Maki–Thompson model for rumour spread is an asymptotically proportional epidemic model with $\alpha = 2$.

It is possible to derive the asymptotic distribution of S_W in the Maki–Thompson model using the technique with Doléans–Dade exponential as in the proof of Theorem 5.5. Since this result has already been given by Watson (1988), we do not give this proof here.

References

ANDERSON, R. M. (1982) *Population Dynamics of Infectious Diseases*. Chapman and Hall, London.

VON BAHR, B. AND MARTIN-LÖF, A. (1980) Threshold limit theorems for some epidemic processes. *Adv. Appl. Prob.* **12,** 319–349.

BAILEY, N. T. J. (1975) *The Mathematical Theory of Infectious Diseases and its Applications*. Griffin, London.

DALEY, D. J. AND KENDALL, D. G. (1965) Stochastic rumours. *J. Inst. Maths. Applics.* **1,** 42–55.

FRAUENTHAL, J. C. (1980) *Mathematical Modelling in Epidemiology*. Springer-Verlag, Berlin.

JACOD, J. AND SHIRYAEV, A. N. (1987) *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin.

KURTZ, T. G. (1981) *Approximation of Population Processes*. SIAM, Philadelpia.

LEFÈVRE, C. (1990) Stochastic epidemic models for S-I-R infectious diseases: a brief survey of the recent general theory. In *Stochastic Processes in Epidemic Theory*, ed. J.-P. Gabriel, C. Lefèvre, and P. Picard. Lecture Notes in Biomathematics **86,** 1–12, Springer-Verlag, Berlin.

MAKI, D. P. AND THOMPSON, M. (1973) *Mathematical Models and Applications*. Prentice-Hall, Englewood Cliffs, NJ.

MARTIN-LÖF, A. (1986) Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J. Appl. Prob.* **23,** 265–282.

PITTEL, B. (1990) On a Daley–Kendall model of random rumours. *J. Appl. Prob.* **27,** 14–27.

WATSON, R. (1981) An application of a martingale central limit theorem to the standard epidemic model. *Stoch. Proc. Appl.* **11,** 79–89.

WATSON, R. (1988) On the size of a rumour. *Stoch. Proc. Appl.* **27,** 141–149.