On the Qualitative Behaviour of SIR Epidemics with Generalized Infection Rate Functions

HAMID EL MAROUFY
ZIAD TAIB

Department of Mathematical Sciences
Division of Mathematical Statistics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg Sweden 2008
On the Qualitative Behaviour of SIR Epidemics
with Generalized Infection Rate Functions

Hamid El Maroufy and Ziad Taib

Department of Mathematical Sciences
Division of Mathematical Statistics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg, Sweden
Göteborg, May 2008
Preprint 2008:20
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2008
On the Qualitative Behavior of SIR Epidemics with Generalized Infection Rate Functions

Hamid El Maroufy\(^a\) and Ziad Taib\(^b\)

\(^{a}\) Department of Applied Mathematics, Beni-Mellal Faculty of Sciences and Technology, B.P : 523, Beni-Mellal, Morocco

\(^{b}\) Department of Mathematics, Chalmers University of Technology and the Göteborg University, S-412 96 Göteborg, Sweden

Abstract

This note considers a homogeneous SIR stochastic epidemic model in which new infection occur at rate \(f_n(x,y)\) where \(x\) and \(y\) are the number of susceptibles at time of infection and \(f_n\) is a positive sequence of real function. Threshold theorems analogous to those of Whittle (1955) and Williams (1971) are fairly proved for this model. Also we examine the shape of the total size distribution for various values of removal rate and suitable value of other important parameters.

Keywords: Epidemic model; Generalized infection rate; Threshold theorems; Total size.

1 Introduction

The purpose of this note is to examine the qualitative properties of stochastic models with generalized infection rate in which the population is divided into three classes of individuals: susceptible, infective and removed individuals. This model can be used e.g. to model the transmission of complex diseases. Mathematically, it is defined as follows. At time \(t\) the \(X(t)\) describes the number susceptibles, \(Y(t)\) the infectives and \(n - X(t) - Y(t)\) the removed with \(X(0) = n\) and \(Y(0) = a\). The epidemic process is thus completely determined by \(\{(X(t), Y(t)); t \geq 0\}\), which is supposed to be a continuous-time Markov chain on the state space:

\[E_{n,a} = \{(x, y), 0 \leq x \leq n, 0 \leq y \leq n + a - x\}\]

with the transition probabilities

\[
\begin{align*}
\Pr\{(X(t + \delta t), Y(t + \delta t)) & = (x - 1, y + 1)(X(t), Y(t)) = (x, y)\} \\
& = f_n(x, y)\delta t + o(\delta t), \\
\Pr\{(X(t + \delta t), Y(t + \delta t)) & = (x, y - 1)(X(t), Y(t)) = (x, y)\} \\
& = \mu y\delta t + o(\delta t)
\end{align*}
\]

(1)

all other transitions having probability \(o(\delta t)\), and the parameter \(\mu\) is known as the removal rate. The process terminates when the number of infectives becomes zero, which will almost surely happen within a finite time.

---

\(^1\)Corresponding author : Tel : 212 63 77 44 03; fax : 212 23 48 52 01; E-mail address : maroufy@fstbm.ac.ma; elmarouf@math.chalmers.se; (H. El Maroufy).
If \( f_n(x, y) = \beta xy \) is chosen, where \( \beta \) is an infection parameter, the model is reduced to the general model (see Bailey (1975, page 88)[1]). There are two threshold theorems for general epidemics, Wittle’s theorem (see, e.g Wittle (1955)[24] and Williams theorem (see, e.g William (1971)[25]), that govern the qualitative behaviour of the epidemic. These two theorems are based respectively on the asymptotic approximation of the distribution of the intensity of the epidemic, \( I = \frac{n - X(\infty)}{n} \), and the total size distribution. These results show that a small change in relative removal rate \( \rho = \frac{\mu}{\beta} \) leads to the qualitative change of the epidemic, and are generalized by Ball and O’Neill (1993)[5] and O’Neill (1995[18],1997[20]) to allow the case when \( f_n(x, y) = \beta \frac{x}{x+y} \). The above theorems all require that the population size approaches infinity, then given small finite size is so large that this limiting result is an acceptable approximation. In this situation N˚ asell (1995)[17], using numerical methods, studied the threshold of the epidemic by illustrating the form of the total size distribution.

If we consider the SIR model with generalized infection rate, we claim that the classical Williams’s theorem and a fairly proof of Wittle’s theorem are not yet obtained. Apart from the paper of Gani and Purdue (1981)[14] that gives an intuitive proof of Wittle’s result, our contribution aims to see how these results can be extended to the model as described by (1). Under some condition on the infection rate \( f_n(x, y) \) we give an algebraic proof of Williams’s theorem in section 2; this proof outlined the explicit formula for the Laplace transforms of the the transitions probabilities obtained by El Maroufy et al (2002)[11]. In the third section we give a rigorous proof of Wittle’s theorem using the coupling method. A qualitative study in the case of small size is examined in the fourth section.

2 Williams’s threshold theorem

In order to establish the counterpart of William’s threshold theorem for our model, we need to restrict the behavior of the infection rate \( f_n(x, y) \). For a sufficiently large \( n \) and a suitable choice of \((x, y)\), \( f_n(x, y) \) should be closer to \( \beta(n)y \) where \( \beta(n) \) is a positive constant that may depend on \( n \). To this end, we now define the class of sequences of functions \( f_n \).

**Definition 1** Let \( \mathcal{L} \) be the set of all real-valued sequences \((x_n, n \geq 0)\) for which there exists \( k \in \mathbb{N} \) such that \( |x_n - n| < k \) for all \( n \in \mathbb{N} \).

**Definition 2** Let \( (f_n)_{n \geq 0} \) be sequences of positive real-valued functions. Then \( (f_n)_{n \geq 0} \in \mathcal{L}_0 \), if, for all \( (x_n)_{n \geq 0} \in \mathcal{L} \),

\[
f_n(x_n, y) \sim \beta(n)y, \text{ when } n \to \infty \text{ and } y \in \mathbb{N}.
\]

Let \( P_{il}(t) = \Pr\{X(t) = i, y(t) = l\} \) : the probability that the epidemic with the state \((n, a)\) at time 0 passes to the state \((i, l)\) at time \( t \) and for \( r = 0,1,\ldots,n \), let \( \Pi_r \) be the
Moreover, if we suppose that \( (\alpha_0, \alpha_1, \ldots, \alpha_r) \) be the non-negative numbers such that
\[
\alpha_0 = a + r - l_r, \quad \alpha_l = l_{w+1} - l_w, \quad \alpha_r = l_r,
\]
with \( l = (l_0, \ldots, l_r) \in \hat{D}_{0a+r} \). Then the set of vectors \( \alpha = (\alpha_0, \ldots, \alpha_r) \) has the same cardinal as the set \( A_r \) defined in Rajarshi (1981)[19]. However, according to Foster (1955)[13], (see also Ball and O’Neill (1993)[5]), the \( A_r \) set is identical to the set of all paths from \((n, a)\) to \((n - r, 0)\), when the epidemic process is viewed as a random walk on \( E_{n,a} \). Then the ballot theorem (Feller (1971)[12]) implies that
\[
|A_r| = c_r^{2r+a-1} \frac{a}{a+r}.
\]
It results from (5) and (6) that
\[
\Pi_r \approx \frac{(2r + a - 1)!a}{r!(r + a)!} \left( \frac{\rho(n)}{\rho(n) + 1} \right)^{a+r} \left( \frac{1}{\rho(n) + 1} \right)^r, \quad r = 0, 1, \ldots
\]  
(7)

With the same algebraic techniques used by Bailey (1975, page 107), it may be seen that the right member of (7) is the \( r \)-th term in the expansion of \( \left( \frac{1 + \rho(n) - |\rho(n)|}{2} \right)^a = (\min\{1, \rho(n)\})^a \), and
\[
\sum_{r=0}^{\infty} \Pi_r = (\min\{1, \rho(n)\})^a,
\]
so the following result is obtained:

**Theorem 4** Suppose that \((f_n)_n \in \mathcal{L}_0\). Then for \(n \) sufficiently large, the probability of a minor epidemic is given by
\[
\Pr\{T < \infty\} \approx (\min\{1, \rho(n)\})^a.
\]  
(8)

If we suppose that \(f_n(x, y) = \beta_n(x,y)\) for \(\alpha \geq 0\) then \((f_n)_n \in \mathcal{L}_0\) then (8) becomes \(\Pr\{T < \infty\} \approx (\min\{1, \rho n^{\alpha-1}\})^a\). In this case, if \(\alpha = 0\) and \(\alpha = 1\), the probability above is the same as that obtained respectively by Rajarshi (1981) [19] for the general epidemic and by Ball and O’Neill (1993) [5] for the modified epidemic.

The limiting distribution given by (7) is similar to that found by Ball and Näsell (1994) [4] and corresponds to the distribution of the final size of birth-death process with the extinction probability given by (8). This interpretation involves that, when \(n\) tends to infinity and \(f_n\) checks the conditions of definition 2, the epidemic process is approached by a birth and death process with birth rate \(\rho(n)\) and death rate \(\rho(n)\) and initial population size \(a\). So a major epidemic can occur with probability \(1 - \rho(n)\) if and only if \(\rho(n) < 1\).

### 3 Whittle’s threshold theorem

In this section we restrict ourselves to the case \(f_n(x, y) = \beta_n(x,y)\). \(x,y\) with \(\beta_n(x,y) = 0\) if \(x\) or \(y\) = 0 where \(\beta_n\) is a specified function that determines the type of infection mechanism, it comprises some infection mechanisms mentioned in epidemic literature. For instance, Clancy (1999a) [8] took \(\beta_n(x,y) = \frac{\beta}{(x+y)^{\alpha}}\); here \(\beta\), as defined in Dietz (1988) [10], is the product of the contact rate and the probability that a successive number of contacts leads to infection. In this case, if \(\alpha = 1\), this gives the model considered by Gleißner (1988) [15], Ball and O’Neill (1993) [5] and Sani et al (2007) [21]. When \(\alpha = 0\), the model is reduced to the general epidemic model. the case \(\alpha = \frac{1}{2}\) was considered by Saunders (1980) [22].

In order to give a rigorous proof of Whittle’s threshold theorem, we begin by defining our model using a construction due to Sellke (1983) [23] (see also Ball (1995) [3] Ball and O’Neill (1999) [6]). Label the initial infectives \(-(a - 1), \ldots, 0\) and the initial susceptibles
1, ..., n. Let \( R_{-(a-1)}, ..., R_0 \) and \( R_1, ..., R_n \) be independent sequences of independent negative exponential random variables with mean \( \mu^{-1} \). For \( j = -(a-1), ..., 0 \) the initial infective remains infectious for a period \( R_j \) and it is then removed, while for \( j = 1, ..., n \) \( R_j \) is the infectious period of the \( j \)-th susceptible to become infected. For \( j = 1, ..., n \) let \( Q_j \) denote the infection tolerance of susceptible \( j \), the \( Q_j \)’s are independent copies of some non-negative exponential random variables having mean 1 and denoted by \( Q(1), ..., Q(n) \) the order statistic associated to \( (Q_j, 1 \leq j \leq n) \).

For \( i = -(a-1), 0, 1, ..., n \), let \( \tau_j \) be the time of individual \( j \)'s infection, with \( \tau_j = 0 \) if \( j = -(a-1), 0 \), and \( \tau_j = +\infty \) if susceptible \( j \) avoids infection. For \( t \geq 0 \), any remaining infective accumulates exposure to infection at rate \( \beta_n(X,Y) \). Our epidemic now proceeds as follows: knowing that \( j \) infections occur before, the \( j+1 \) susceptible becomes infected when its total exposure to infection (see Ball and O’Neill (1993[5], 1999[6]))

\[
\chi_j(t) = \int_0^{t_j} \beta_n(X(u), Y(u))Y(u)du, \quad \text{with} \quad t_j = \min \left\{ t, \max_{-(a-1) \leq i \leq j} (\tau_i + R_i) \right\} \tag{10}
\]

reaches \( Q_{j+1} \). The epidemic ceases as soon as no more infectives are left in the population. With these arguments, the final size of the epidemic is equal to

\[
T = \min \{ r \in [0, n] : Q_{(r+1)} > \chi_r(\infty) \} \tag{11}
\]

thus

\[
\{ T \geq k \} = \bigcap_{r=1}^{k} \{ Q_r \leq \chi_{r-1}(\infty) \}, \quad \forall k \in [0, N]. \tag{12}
\]

By using the intensity of epidemic defined in section 1 we have, for \( t \geq 0 \),

\[
n(1 - I) \leq X(t) \leq n \quad \text{and} \quad 0 \leq X(t) + Y(t) \leq n + a.
\]

Let \( A_I = \{(x,y) ; n(1 - I) \leq x \leq n \quad \text{and} \quad 0 \leq x + y \leq n + a \} \) and let \( m_I \) and \( M \) be respectively suitable lower and upper bounds of \( \beta_n(x,y) \) and \( \beta_n(x,y) \) over the set \( A_I \) and \( E_{n,a} \) respectively. Then for \( t \geq 0 \) we obtain for \( (X(t), Y(t)) \in A_I \)

\[
n(1 - I)m_IY(t) \leq \beta_n(X(t), Y(t))X(t)Y(t) \leq MY(t) \tag{13}
\]

These inequalities show that the process can be sandwiched between two other epidemic processes each having removal rate \( \mu \). The first is slow and the second is fast with the total exposure to infection, such that \( j \) infections occur before, respectively equal to

\[
\hat{\chi}^I_j(t) = (1 - I)m_I \int_0^{t_j} \frac{Y(u)}{X(u)}du \tag{14}
\]

and

\[
\hat{\chi}_j(t) = M \int_0^{t_j} \frac{Y(u)}{X(u)}du \tag{15}
\]

From (10), (14) and (15), and using (13), we deduce that:

\[
\hat{\chi}^I_j(\infty) \leq \chi_j(\infty) \leq \hat{\chi}_j(\infty). \tag{16}
\]
Let 
\[ \tilde{T}^I = \min\{r \in \{0, \ldots, n\} : Q(r) > \tilde{\chi}^I(\infty) \} \]
and 
\[ \tilde{T} = \min\{r \in \{0, \ldots, n\} : Q(r) > \tilde{\chi}(\infty) \} \]
the final sizes of the two epidemics respectively. Then, using (12) and applying (16), we find that

\[ \{\tilde{T} \leq k\} \subseteq \{T \leq K\} \subseteq \{\tilde{T}^I \leq K\} \quad \forall k \in \mathbb{N} \tag{17} \]

the second inclusion of (17) implies, for \(i \in \]0, 1[, that

\[ \Pr\{T \leq ni\} = \Pr\{T \leq ni, \tilde{T}^I \leq ni\} = \Pr\{I \leq i, \tilde{T}^I \leq ni\} \leq \Pr\{I \leq i, \tilde{T} \leq ni\} \leq \Pr\{\tilde{T} \leq ni\} \tag{18} \]

where \(\tilde{T}\) is the final size of the epidemic with infection rate \(n(1 - i)m_i\).

Combining (17) and (18), we obtain

\[ \Pr\{\tilde{T} \leq ni\} \leq \Pr\{T \leq ni\} \leq \Pr\{\tilde{T}^i \leq ni\} \]

on the other hand \(\Pr\{\tilde{T} \leq ni\} = \Pr\{\tilde{T} < \infty\} - \Pr\{ni < \tilde{T} < \infty\}\) and, when \(n\) is sufficiently large, \(\Pr\{ni < \tilde{T} < \infty\} \approx 0\). Moreover, \(\Pr\{\tilde{T}^i \leq ni\} \leq \Pr\{\tilde{T}^i < \infty\}\). Consequently, by considering the following distribution \(\pi_i = \Pr\{I \leq i\} = \Pr\{T \leq ni\} = \sum_{r=0}^{ni} \Pi_r\) and using Theorem 4, the following result is obtained:

**Theorem 5** For sufficiently large \(n\)

\[ \left(\min\left\{\frac{\mu}{M}, 1\right\}\right)^a \leq \pi_i \leq \left(\min\left\{\frac{\mu}{n(1 - i)m_i}, 1\right\}\right)^a \tag{20} \]

The statement in (20) constitutes Whittle’s stochastic threshold theorem. It may be interpreted by saying that if \(\mu > M\) then \(\pi_i = 1\), so there is zero probability of an epidemic exceeding any intensity \(i \in ]0, 1[\).

In a particular case as \(\beta_n(X(t), Y(t)) = \frac{\beta^{n-1}_{1-a}}{(X(t)+Y(t))^\alpha}\) where \(\alpha\) is defined as previously, we obtain

\[ m_i = \frac{\beta^{n-1}_{1-a}}{(n+a)^\alpha} \quad \text{and} \quad M = \beta. \]

The statement in (20) becomes, for sufficiently large \(n\):

\[ \left(\min\{\rho, 1\}\right)^a \leq \pi_i \leq \left(\min\left\{\frac{\rho(n + a)^\alpha}{n^\alpha(1 - i)}, 1\right\}\right)^a \tag{21} \]

We see that, in this case, the probability that an epidemic exceeds the size \(ni\) is approximately \(1 - \rho^a\). When \(\alpha = 1\), the upper boundary in (20), is greater than \(\left(\min\{1, \frac{\rho}{1 - i}\}\right)^a\), which is obtained by Gani and Purdue (1984) and Ball and O’Neill (1993) then the methods outlined here give approximation bounds of \(\pi_i\) for large \(n\).

### 4 The Shape of the total size distribution

In this section we are concerned with the shape of the distribution curve of the total size considering the following particular infection rate

\[ f_n(x, y) = \frac{\beta xy}{(x + y)^\alpha}, \quad \alpha \geq 0, \tag{22} \]
where $\beta$ and $\alpha$ are as previously defined. Bailey(1975), Ball and O’Neill(1993)[5], Nåsell (1995)[17] and Clancy(1999)[9] give, for an epidemic model with particular infection rate functions, the total size distribution for various values of the removed rate $\rho$, remarking that the distribution curve can for a small number of initial infectives take one of the two shapes called J-shape and U-shape. The J-shaped curve can be interpreted as describing a minor epidemic, while a U-shaped curve is associated to a minor or major epidemic. More extensive results are shown in Figure 1-3 over a suitable range of $\alpha$.

Figure 1: Distribution of the final sizes, $n = 1000$, $\rho = 600$ (representing general model), $\rho_{1/2} = 54.7$ (Saunders’s model), $\rho_1 = 5$ (modified model), $\rho_2 = 0.041$ (our proposed model).

Figure 2: Distribution of the final sizes, $n = 1000$, $\rho = 120$ (representing general model), $\rho_{1/2} = 11$ (Saouders’s model), $\rho_1 = 1$ (modified model), $\rho_2 = 0.008$ (our proposed model)
Figure 3: Distribution of the final sizes, \( n = 1000 \), \( \rho_0 = 60 \) (representing general model), \( \rho_1 = 0.5 \) (modified model) \( \rho_2 = 0.0041 \). (our proposed model)

All the probabilities used here were originally calculated using the following two-dimensional recursive equation of the Laplace-transform \( g_{i,l} = \lim_{v \to 0} \hat{P}_{il}(0) \), where the quantities \( \hat{P}_{il}(0) \) for \( i = 0, 1, \ldots, n \) and \( l = 1, \ldots, n+a-i \) given by (4) verify the following recursive equations

\[
g_{n,l} = \mu^{a-l} \frac{a!}{l!} \prod_{k=l}^{a} \left( f_n(n, k) + \mu k \right)^{-1}, \quad l = 1, \ldots, a
\]

and

\[
g_{i,l} = \sum_{\text{max}(2,l) \leq h \leq n+a-i} \mu^{h-l} \frac{h!}{l!} f_n(i+1, h-1) \frac{g_{i+1,h-1}}{\prod_{k=l}^{h} (\mu k + f_n(i, k))}, \quad l = 1, \ldots, n+a-i.
\]

The above two-dimensional system is perfectly adequate for computing purposes. Since the threshold behavior of the four epidemics is controlled respectively by \( \frac{\rho_0}{n}, \frac{\rho_1}{\sqrt{n}}, \rho_1 \) and \( \rho_2 n \), then if we set \( \rho_0 = n^{1-\alpha} \rho_1 \), \( \rho_\alpha \) will be above its threshold value \( \rho_\alpha = n^{1-\alpha} \) if and only if \( \rho_1 \) is above its threshold \( \rho_1 = 1 \). Under this condition, the curves of the total size distribution have the same shape for all four models.

It is clear that, from Figure 1, when \( \rho_\alpha = n^{1-\alpha}, \alpha = 0, 1/2, 1, 2 \) is above its threshold then all curves fall rapidly and tend to be null as the total increases. In other words, the curve is J-shaped (Figure 1); this illustrates the fact that the epidemic dies out quickly and becomes minor. However, on one hand, when the relative rates \( \frac{\rho_0}{n}, \frac{\rho_1}{\sqrt{n}}, \rho_1 \) and \( \rho_2 n \) are respectively below their thresholds \( n, n^{1/2}, 1 \) and \( n^{-1} \) then the curves are U-shaped (Figure 3) but not more pronounced (Figure 2, \( a = 10 \)). On the other hand, Figure 3 illustrates the fact that the epidemic is major with increasing the degree \( \alpha \), in the sense that there is higher probability of none of the initial susceptibles contracting the disease in the general epidemic and Saunderson’s epidemic than in others.
5 Conclusions

In this note we have examined the qualitative properties described for an SIR epidemic model with a generalized infection mechanism. We may obtain the same result by considering a more generalized removal rate $\mu_n(x_n, y) \sim \mu(n)y$ when $n$ is sufficiently large for all sequences $(x_n)_{n\geq0} \in \mathcal{E}$. As illustrated in sections 2 and 3, the method used to prove rigourously Wiliam’s and Wittle’s threshold theorems is versatile, and can be adapted to various multipopulation SIR epidemic models. This will be investigated in future research.

Acknowledgments

This work has been done while H. El Maroufy was visiting the Department of Mathematical Sciences at Chalmers University of Technology and the Göteborg University. He would like to thank the Faculty of that Department. His visit to Sweden was made possible through a grant from the Swedish Institute to which he is very grateful.

References


