LAW OF LARGE NUMBERS FOR THE SIRS EPIDEMIC

R. G. DOLGOARSHINNYKH

ABSTRACT. We establish law of large numbers for SIRS stochastic epidemic processes: as the population size increases the paths of SIRS epidemic processes converge to the trajectories of a two dimensional dynamical system. If the reproductive number, R_0, is greater than one then the limiting system has an interior fixed point so that the epidemic processes started with a nontrivial proportion of initially infected are endemic. We show global stability of the interior fixed point of the limiting dynamical system.

1. INTRODUCTION

The SIRS epidemic models were introduced in 1933 by Kermack and McKendrick (5) to describe endemic infections. In any time interval $(t, t+\Delta t)$ a susceptible individual is equally likely to come into contact with any of the currently infected individuals and the average number of such contacts is proportional to a fixed fraction of currently infected individuals, $\theta_1 I_t/N$. Susceptible individuals are equally susceptible to the disease and the probability of infection upon a contact with an infected individual is θ_2. In the same time interval, any infected individual may recover and become temporarily immune at rate ρ and an immune individual may lose his immunity and become susceptible at rate 1. That is, the unit of time is taken to be equal to the average immune period. The population is assumed closed so that there is no inflow or outflow of individuals.

Let $N \in \mathbb{N}$ be the population size parameter. For $t \geq 0$, let

$$S_t = \text{the number of susceptible individuals at time } t,$$
$$I_t = \text{the number of infected individuals at time } t,$$
$$R_t = \text{the number of recovered (immune) individuals at time } t,$$

and let $s_t = S_t/N$, $i_t = I_t/N$ and $r_t = R_t/N$. Since $S_t + I_t + R_t \equiv N$, the pair (S_t, I_t) or (s_t, i_t) completely describes the state of the system at any time t.

We say that a transition occurs whenever an individual changes state. Under the assumptions, the process (S_t, I_t) is a pure jump Markov process

Key words and phrases. Stochastic epidemic model, SIRS epidemic processes, Law of large numbers, Endemic infection, Threshold phenomena in epidemic processes.

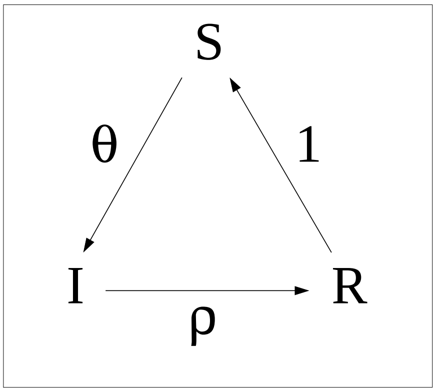

Figure 1. SIRS model.

The letters S, I, and R refer to the possible states an individual can assume in turn — susceptible, infected and recovered (immune). The individual transition rates are shown above the arrows.

---

with the following instantaneous transition rates

$$\begin{aligned} q\{(S, I) \to (S-1, I+1)\} &= \theta SI/N, \\ q\{(S, I) \to (S, I-1)\} &= \rho I, \\ q\{(S, I) \to (S+1, I)\} &= R, \end{aligned}$$

where $\theta = \theta_1 \times \theta_2$. It will often be more convenient to work with the scaled process $\gamma_t^N = \gamma_t = (s_t, i_t)$. This is a continuous time Markov Chain with state space $K^N = \{(s, i) : (Ns, Ni) \in \mathbb{Z}_+^2, s + i \leq 1\}$ and transition rates

$$\begin{aligned} q\{(s, i) \to (s - N^{-1}, i + N^{-1})\} &= N\theta si, \\ q\{(s, i) \to (s, i - N^{-1})\} &= N\rho i, \\ q\{(s, i) \to (s + N^{-1}, i)\} &= Nr. \end{aligned} \tag{1}$$

The transition rules (1) imply that for small $h > 0$

$$\begin{aligned} \mathrm{E}(s_{t+h}|\mathcal{F}_t) &= s_t + r_t h - \theta i_t s_t h + o(h), \\ \mathrm{E}(i_{t+h}|\mathcal{F}_t) &= i_t + \theta i_t s_t h - \rho i_t h + o(h), \end{aligned}$$

where $\mathcal{F}_t$ contains all events that happen before or at time t. Letting $h \to 0$ suggests that the random SIRS paths (at least over short periods of time) may be approximated in mean by the solutions of the system of ordinary differential equations

$$\begin{cases} \dfrac{ds_t}{dt} = r_t - \theta i_t s_t, \\ \dfrac{di_t}{dt} = \theta i_t s_t - \rho i_t. \end{cases} \tag{2}$$

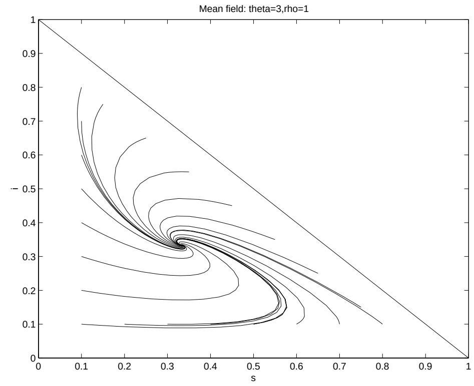

FIGURE 2. Mean field.

Trajectories of (2) for $\theta > \rho$ with started at initial points with $i_0 > 0$.

We will show in sec. 3 that the random SIRS paths, γ^N, tend to solutions of system (2) as the populations size N tends to infinity. This is a first order approximation to the random paths, or the strong law of large numbers for the problem. Here and later we will often refer to the random SIRS paths by γ^N or γ and to the solutions of the mean path ODE by $\bar{\gamma}$. Trajectories of the dynamical system (2) lie in $K = \{(s,i) : s, i \geq 0,\, s+i \leq 1\}$.

The system has fixed points, that is points such that the rates of change in s and i at these points are all equal to 0. Solving for such points, we get two possibilities; either

$$s = \frac{\rho}{\theta} \equiv s_\infty \quad \text{and} \quad i = \frac{\theta - \rho}{\theta(1+\rho)} = \frac{1 - \rho/\theta}{1+\rho} \equiv i_\infty$$

or $s = 1$ and $i = 0$. When $\rho < \theta$ the fixed point $\gamma_\infty = (s_\infty, i_\infty)$ is in the interior of the triangle K; when $\rho \geq \theta$ the only fixed point of the dynamical system in K is $s = 1$, $i = 0$. We will show in sec. 5 that the solutions of (2) started at initial points with $i > 0$ converge to γ_∞ when $\rho < \theta$ and to $(1, 0)$ if $\rho \geq \theta$ as t tends to infinity.

Once we establish that the random paths γ^N converge to solutions $\bar{\gamma}$ of (2) we can conclude that if the rate of infection θ is less than or equal to the rate of recovery ρ, the infection tends to become rapidly extinct by following the mean path to the point where $s = 1$ and $i = 0$. On the other hand, if ρ is less than θ then with a significant fraction of the population initially infected we expect the epidemic to follow the mean path to the fixed point in the interior of K and become endemic; that is, the fraction of infected individuals in the population remains significant for a long time. In the case when $\rho < \theta$, the state of the system turns out to be well approximated by the endemic level, γ_∞, for a long time.

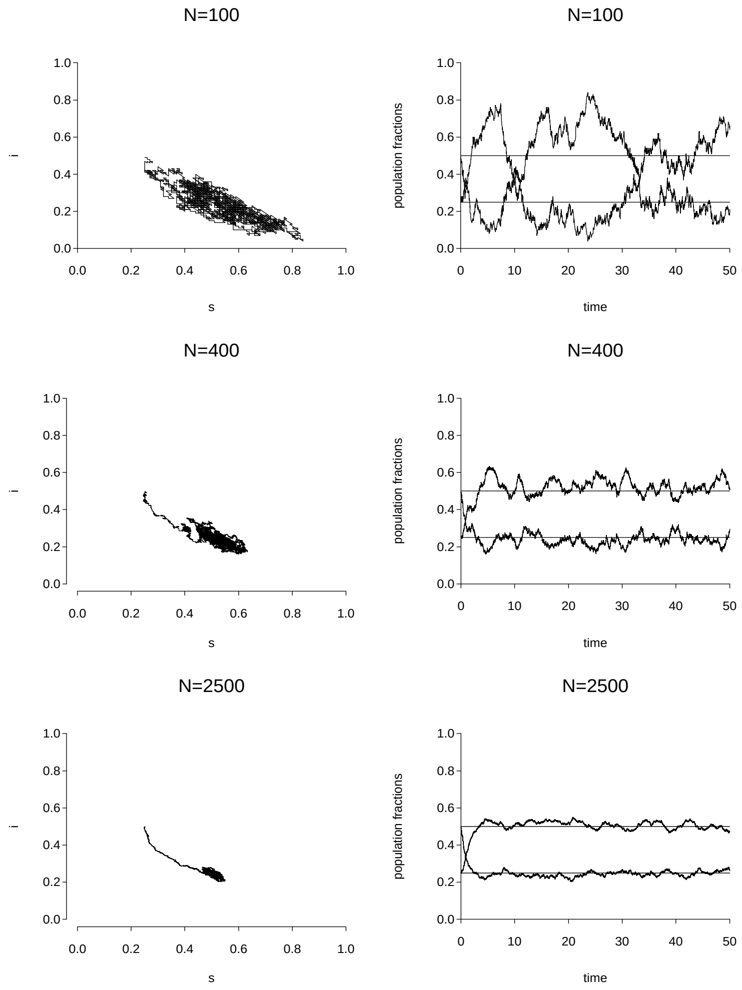

FIGURE 3. Initial infection spread.

The simulations are for population sizes $N = 100, 400, 2500$ and for parameters $\theta = 2$, $\rho = 1$. For this choice of parameters the fixed point is $\gamma_\infty = (s_\infty, i_\infty) = (1/2, 1/4)$. All three simulations were started at $(s_0, i_0) = (1/4, 1/2) \times N$. The left plots show the dynamics of the infection spread in terms of fractions, s and i, of susceptibles and infecteds for the three population sizes over time period of length 50. As N increases the random paths are seen to follow the limiting deterministic paths more closely. In the plots on the right side values of s and i are plotted over time. We see that the fluctuations around the fixed point γ_∞ decrease with N.

The existence of a threshold at the level $\rho = \theta$, that is, qualitatively different behavior of the model for $\rho < \theta$ and $\rho \geq \theta$, is not unexpected. The reproduction number, R_0, is defined as the (average) number of secondary cases that are produced by one infected individual in wholly susceptible population. It has been established that large epidemics can occur when $R_0 > 1$, and the infection quickly becomes extinct when $R_0 < 1$ (see e.g. (6)). For the SIRS model with transitions given by (1) the reproduction number $R_0 = \theta \times 1/\rho = \theta/\rho$. Therefore, the epidemics die out quickly when $R_0 < 1$ or $\theta < \rho$ and but not when $R_0 > 1$, that is when $\theta > \rho$. The case $R_0 = 1$ for the SIRS model also leads to extinction of the infection after a short period of time.

2. SIRS Processes as Time Changed Poisson Processes

Before we proceed with the proof of our main theorem we construct a version of the random SIRS processes γ^N by applying random time changes to a collection of Poisson processes see e.g. (3) chap 4, §4. In particular, let Y_1, Y_2, Y_3 be independent rate one Poisson processes and let $y_k(\cdot) = N^{-1}Y_k(N\cdot)$ for $k = 1, 2, 3$ and $N \in \mathbb{N}$. If the initial population fractions $\gamma_0 = (s_0,\, i_0)$ are nonrandom then there is a version of γ satisfying

$$\text{(3)} \qquad \begin{cases} s_t = s_0 - y_1\left(\displaystyle\int_0^t \theta s_u i_u \, du\right) + y_3\left(\displaystyle\int_0^t r_u \, du\right) \\ i_t = i_0 + y_1\left(\displaystyle\int_0^t \theta s_u i_u \, du\right) - y_2\left(\displaystyle\int_0^t \rho i_u \, du\right). \end{cases}$$

In what follows it will be useful to define

$$\text{(4)} \qquad \beta^1(\gamma) = \theta s i, \quad \beta^2(\gamma) = \rho i, \quad \beta^3(\gamma) = r = 1 - s - i.$$

Note that for any $t > 0$ and $k = 1, 2, 3$

$$\int_0^t \beta^k(\gamma_u)\, du \leq (\theta + \rho + 1)t < \infty$$

and therefore the $Y_k(Nt)$'s are well defined for all $t \in [0, \infty)$.

For random initial conditions we can construct the processes γ^N satisfying (3) by first generating initial conditions independently of Y_1, Y_2, Y_3.

3. Proof of the Law of Large Numbers

Fix $T > 0$. Let $\gamma = \gamma^N$ be a sequence of SIRS processes, indexed by the population size parameter N, satisfying (3) for $t \leq T$.

Theorem 1. *If* $\lim_{N\to\infty} \gamma_0 = \bar{\gamma}_0$ *and* $\bar{\gamma}$ *is a solution of (2) on* $[0, T]$ *with an initial condition* $\bar{\gamma}_0$ *then*

$$\lim_{N\to\infty} \sup_{t \leq T} |\gamma - \bar{\gamma}| = 0 \quad a.s.$$

Proof. Since γ satisfies (3) for $t \leq T$,

$$|s_t - \bar{s}_t| \leq |s_0 - \bar{s}_0| + \left|\tilde{y}_1\left(\int_0^t \beta^1(\gamma_u)\,du\right)\right| + \left|\tilde{y}_3\left(\int_0^t \beta^3(\gamma_u)\,du\right)\right| + \left|\int_0^t \beta^1(\gamma_u)\,du - \beta^1(\bar{\gamma}_u)\,du\right| + \left|\int_0^t \beta^3(\gamma_u)\,du - \beta^3(\bar{\gamma}_u)\,du\right|$$

where $\tilde{y}_k(s) = y_k(s) - s$ are centered and scaled Poisson processes and β^k's are defined by (4). Let $\epsilon_k(t) = \left|\tilde{y}_k\left(\int_0^t \beta^k(\gamma_u)\,du\right)\right|$. From the strong law of large numbers for a Poisson process, we have for $0 \leq v < \infty$ and $k = 1, 2, 3$

$$\lim_{N\to\infty} \sup_{u\leq v} |\tilde{y}_k(u)| = 0 \quad \text{a.s.}$$

Since $\beta^k(\gamma_u) \leq (\theta + \rho + 1)$ it follows that

$$\lim_{N\to\infty} \sup_{t\leq T} \epsilon_k(t) = 0 \quad \text{a.s.}$$

for $k = 1, 2, 3$.

Because the functions β^k are Lipschitz continuous as functions of γ, there exists a constant $M > 0$ such that

$$|s_t - \bar{s}_t| \leq |s_0 - \bar{s}_0| + \epsilon_1(t) + \epsilon_3(t) + \int_0^t M|\gamma_u - \bar{\gamma}_u|\,du.$$

The same argument leads to a similar bound on $|i_t - \bar{\imath}_t|$ and combining the two we get

$$|\gamma_t - \bar{\gamma}_t| \leq 2\left(|\gamma_0 - \bar{\gamma}_0| + \epsilon_1(t) + \epsilon_2(t) + \epsilon_3(t) + \int_0^t M|\gamma_u - \bar{\gamma}_u|\,du\right) \leq 2\left(|\gamma_0 - \bar{\gamma}_0| + \sup_k \sup_{t\leq T} \epsilon_k(t) + \int_0^t M|\gamma_u - \bar{\gamma}_u|\,du\right).$$

Hence by Gronwall's inequality stated below (for proof see e.g. (4), chap 3)

$$|\gamma_t - \bar{\gamma}_t| \leq 2\big(|\gamma_0 - \bar{\gamma}_0| + \sup_k \sup_{t\leq T} \epsilon_k(t)\big)\, e^{2MT}$$

and the conclusion of the theorem follows. □

Gronwall's Inequality.

Let f be an integrable function on $[0, T]$. If $M > 0$ and

$$0 \leq f(t) \leq \epsilon + M \int_0^t f(s)\,ds, \quad \text{for } 0 \leq t \leq T,$$

then

$$f(t) \leq \epsilon e^{Mt}, \quad \text{for } 0 \leq t \leq T.$$

4. Construction of Lyapunov Function

Let $f(x)$ be a continuous function of x defined on an open set containing 0. Let $x : [0, \infty) \to \mathbb{R}^d$ solve $\dot{x}_t = f(x_t)$. A function $V(x)$ defined in a neighborhood of $x = 0$ is a Lyapunov function if it has continuous partial derivatives, $V(x) \geq 0$ and the trajectory derivative of V satisfies

$$\dot{V}(x_t) = (\operatorname{grad} V) \cdot f(x_t) \leq 0,$$

where the dot denotes scalar multiplication and

$$\operatorname{grad} V = (\partial V/\partial x^1, \ldots, \partial V/\partial x^d).$$

Constructing a Lyapunov function allows us to show global asymptotic stability of trajectories of (2) started at initial points such that $i > 0$; that is, if $\bar{\gamma}$ is a solution of (2) such that $i_0 > 0$ then $\lim_{t\to\infty} \bar{\gamma}_t = \gamma_\infty$.

To construct a Lyapunov function for the mean path vector field we need to find a function that is positive on the domain and has a negative derivative along the directions of the vector field. We will find it convenient to switch to (i, r) coordinates. In these coordinates our vector field becomes

$$\begin{cases} \dfrac{di}{dt} = \theta(1 - i_t - r_t)i_t - \rho i_t \\ \dfrac{dr}{dt} = \rho i_t - r_t. \end{cases}$$

Let $x_t = i_t - i_\infty$, $y_t = r_t - r_\infty$. Showing that $(0, 0)$ is globally asymptotically stable in the (x, y) coordinates is equivalent to showing that γ_∞ is globally asymptotically stable in the old coordinates. In (x, y) coordinates,

$$\begin{cases} \dfrac{dx}{dt} = \theta\big(1 - (x_t + i_\infty) - (y_t + r_\infty)\big)(x_t + i_\infty) - \rho(x_t + i_\infty) \\ \dfrac{dy}{dt} = \rho(x_t + i_\infty) - (y_t + r_\infty) \end{cases}$$

and simplifying we get

$$\begin{cases} \dfrac{dx}{dt} = -\theta x_t^2 - \theta x_t y_t - \dfrac{\theta - \rho}{1 + \rho} x_t - \dfrac{\theta - \rho}{1 + \rho} y_t \\ \dfrac{dy}{dt} = \rho x_t - y_t. \end{cases} \tag{5}$$

Following the prescription that can be found in e.g. (2) we seek a Lyapunov function of the form

$$v(x, y) = F(x) + G(y).$$

Then

$$\frac{dv(x_t, y_t)}{dt} = \frac{dF}{dx} \times \left(-\theta x_t^2 - \theta x_t y_t - \frac{\theta - \rho}{1 + \rho} x_t - \frac{\theta - \rho}{1 + \rho} y_t \right) + \frac{dG}{dy} \times (\rho x_t - y_t).$$

We further restrict our search to functions v such that the cross terms disappear, that is (omitting index t to simplify expressions)

$$\frac{dF}{dx} \times \left(-\theta xy - \frac{\theta-\rho}{1+\rho}\,y\right) + \frac{dG}{dy} \times \rho x \equiv 0;$$

implying that

$$\begin{aligned}\frac{dF}{dx} \times y\left(\theta x + \frac{\theta-\rho}{1+\rho}\right) &= \frac{dG}{dy} \times \rho x \\ \frac{dF}{dx} \times \left(\theta x + \frac{\theta-\rho}{1+\rho}\right)\frac{1}{\rho x} &= \frac{dG}{dy} \times \frac{1}{y} \\ \frac{dF}{dx} \times \frac{x+i_\infty}{x} &= \frac{dG}{dy} \times \frac{s_\infty}{y}.\end{aligned}$$

This suggests the possibility

$$\frac{dF}{dx} = \frac{x}{x+i_\infty}, \quad \frac{dG}{dy} = \frac{y}{s_\infty},$$

which when integrated gives

$$F(x) = x - i_\infty \log\left(1 + \frac{x}{i_\infty}\right), \quad G(y) = \frac{y^2}{2s_\infty}$$

and

$$v(x,y) = x - i_\infty \log\left(1+\frac{x}{i_\infty}\right) + \frac{y^2}{2s_\infty}. \tag{6}$$

It is not hard to see that $v(x,y) \geq 0$ on the domain of interest, namely on

$$\{(x,y)\,|\,x > -i_\infty,\, y \geq -r_\infty,\, x+y \leq s_\infty\}$$

and

$$\begin{aligned}\frac{dv}{dt} &= \frac{dF}{dx} \times \left(-\theta x^2 - \frac{\theta-\rho}{1+\rho}\,x\right) + \frac{dG}{dy} \times (-y) \\ &= -\frac{x}{x+i_\infty}\,\theta x(x+i_\infty) - y^2/s_\infty = -\theta\,(x^2 + y^2/\rho) \leq 0.\end{aligned}$$

5. Stability of Interior Fixed Points

In this section we examine stability of the interior fixed point, γ_∞, for the limiting deterministic dynamical system when $\rho < \theta$. We show that for any initial condition with $i_0 > 0$ and hence $x_0 = i_0 - i_\infty > -i_\infty$

$$\lim_{t\to\infty} \gamma_t = \gamma_\infty.$$

Recall that $x_t = i_t - i_\infty$ and $y_t = r_t - r_\infty$ and it is enough to show that the system (5) is globally asymptotically stable. In this section, we continue to work in (x,y) coordinates.

Lemma 1. *Let (x_t, y_t) be the solution of (5) started at (x_0, y_0), where $x_0 > -i_\infty$. There exists $x_* > -i_\infty$ such that $x_t > x_*$ for all $t \geq 0$.*

Proof. Since $v(x, 0) \uparrow +\infty$ as $x \downarrow -i_\infty$ there exists an $x_* < x$ such that $v(x_*, 0) > v(x_0, y_0)$ and $x_* > -i_\infty$. Suppose that the solution of (2) started at (x_0, y_0) intersects level $x = x_*$ at some point (x_*, y_*). Let v be the Lyapunov function defined by (6). Since v is nonincreasing along the trajectories of (5),

$$v(x_*, y_*) \leq v(x_0, y_0),$$

but by construction

$$v(x_*, y_*) \geq v(x_*, 0) > v(x_0, y_0)$$

and we arrive at a contradiction. Therefore $x_t > x_*$ for all $t \geq 0$. □

Lemma 2. *Any solution of (5) started at a point (x_0, y_0) such that $x_0 > -i_\infty$ is asymptotically stable.*

Proof. The linearization of (5) near the point (0,0) is

$$\begin{cases} \dfrac{dx}{dt} = -\dfrac{\theta - \rho}{1 + \rho} x - \dfrac{\theta - \rho}{1 + \rho} y \\ \dfrac{dy}{dt} = \rho x - y \end{cases} := A \times (x, y)^T$$

and A has eigenvalues

$$\lambda_{1,2} = \frac{1}{2} \left(-\frac{1 + \theta}{1 + \rho} \pm \sqrt{\left(\frac{1 + \theta}{1 + \rho} \right)^2 - 4 \frac{\rho(\theta - \rho)}{1 + \rho}} \right),$$

with negative real parts, since $\theta > \rho > 0$. Therefore, system (5) is *Lyapunov stable*, that is for any $\varepsilon > 0$ there exists a $\delta > 0$ such that a solution started in a δ-neighborhood, U_δ, of the stable point never leaves the ε-neighborhood of the stable point, U_ε, see e.g. (1) chap. 3.

Let $\varepsilon > 0$ and δ be as above and x_* be as in Lemma 1. Suppose the trajectory started at (x_0, y_0) never enters U_δ. Then the trajectory will lie in $K_* - U_\delta$, where $K_* = \{(x, y) \,|\, x > x_*,\, y \geq -r_\infty,\, x + y \leq s_\infty\}$. But

$$\frac{dv}{dt} = -\theta \, (x^2 + y^2/\rho) < -C$$

in $K_* - U_\delta$ for some constant $C > 0$. So that

$$v(x_t, y_t) = v(x_0, y_0) + \int_0^t v_t' \, dt \leq v(x_0, y_0) - Ct.$$

If $t \to \infty$, v will eventually have to become negative and we arrive at a contradiction. Therefore the trajectory falls into U_δ eventually and hence $(x_t, y_t) \in U_\epsilon$ for all t after that happens. Since ϵ was arbitrary the trajectory tends to the stable point as t goes to infinity. □

References

[1] V. Arnold (1992) *Ordinary Differential Equations.* Springer, New York

[2] Barabashin (1970) *Introduction to the Theory of Stability.* Wolters-Noordhoff, Netherlands

[3] S. Ethier and T. Kurtz (1985) *Markov Processes. Characterization and Convergence.* Wiley & Sons, NY

[4] P. Hartman (1964) *Ordinary Differential Equations.* Wiley & Sons, NY

[5] W. Kermack and A. McKendrick(1933) Contributions to the mathematical theory of epidemics, III - Further studies of the problem of endemicity. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematics.* **141** 94–122.

[6] D. Mollison, ed. (1995) *Epidemic Models: Their Structure and Relation to Data.* Cambridge, Cambridge