1 Introduction to Epidemic Modelling

1.1 Some Background

Infectious agents have had decisive influences on the history of mankind. Fourteenth century Black Death has taken lives of about a third of Europe's population at the time. The first major epidemic in the USA was Yellow Fever epidemic in Philadelphia in 1793, in which 5,000 people died out of a population of 50,000. This epidemic has had a major impact on the life and politics of the country. Thucydides describes the Plague of Athens (430-428 BC): 1,050 of 4,000 soldiers on an expedition died of a disease. Thucydides gives a detailed account of symptoms: some so horrendous that the last one - amnesia - seems a blessing (Bailey, 1975). An interesting feature of this account is that there is no mention of person-to-person contagion, which we now suspect with most new diseases. It was not until the 19th century that the person-to-person contageon was beginning to be discussed. In this course, we will mostly be interested in modelling infectious diseases, where the major means of disease spread comes from the person-to-person interaction.

The practical use of epidemic models must rely heavily on the realism put into the models. This doesn't mean that a reasonable model can include all possible effects but rather incorporate the mechanisms in the simplest possible fashion so as to maintain major components that influence disease propagation. Great care should be taken before epidemic models are used for prediction of real phenomena. However, even simple models should, and often do, pose important questions about the underlying mechanisms of infection spread and possible means of control of the disease or epidemic.

We begin with classical papers by Kermack and McKendrick (1927, 1932, and 1933). These papers have had a major influence on the development of mathematical models for disease spread and are still relevant in many epidemic situations. The first of these papers laid out a foundation for modelling infections which, after recovery, confer complete immunity (or in case of lethal diseases - death). The population is taken to be constant - no births or deaths other than from the disease are possible - consistent with the course of an epidemic being short compared with the life time of an individual. If

a group of infected individuals is introduced into a large population, a basic problem is to describe the spread of the infection within the population as a function of time. In the course of time the epidemic may come to an end. One of the most important questions in epidemiology is to ascertain whether this occurs only when all of the initially susceptible individuals have contracted the disease or if some interplay of infectivity, recovery, and mortality factors may result in epidemic "die out" with many susceptibles still present in the unaffected population.

In their first paper Kermack and McKendrick (1927) start with the assumption that all members of the community are initially equally susceptible to the disease, and that a complete immunity is conferred after the infection. The population is divided into three distinct classes: the susceptibles, S, - healthy individuals who can catch the disease; the infecteds, I, - those who have the disease and can transmit it; and the removed, R, - individuals who have had the disease and are now immune to the infection (or removed from further propagation of the disease by some other means). Schematically, the individual goes through consecutive states $S \to I \to R$. Such models are often called the SIR models.

1.2 General Epidemic Process

A particular instance of the SIR model is the *general epidemic process* (Kermack and McKendrick, 1927). Let S_t, I_t, and R_t be the number of susceptible, infected and removed individuals, respectively, at time t. Assume that

- $S_t + I_t + R_t \equiv N$ (i.e. the population is closed);
- an individual comes into contact with any another individual at the rate α_1 per unit time;
- upon contact with an infected a susceptible individual contracts the disease with probability α_2, at which time he immediately becomes infected and infectious (no incubation period);
- infecteds recover at an individual rate ρ per unit time.

This defines a continuous time Markov Chain with the state (S_t, I_t). Conditional on $S_t = S$ and $I_t = I$

$$\begin{aligned} \mathrm{P}_t(S_{t+h} = S-1, I_{t+h} = I+1) &= \alpha SIh + \mathrm{o}(h) \\ \mathrm{P}_t(S_{t+h} = S, I_{t+h} = I-1) &= \rho Ih + \mathrm{o}(h), \end{aligned}$$

where $\alpha = \alpha_1 \times \alpha_2$, and

$$\begin{aligned} \mathrm{E}_t(S_{t+h} - S_t) &= -\alpha SIh + \mathrm{o}(h) \\ \mathrm{E}_t(I_{t+h} - I_t) &= \alpha SIh - \rho Ih + \mathrm{o}(h). \end{aligned}$$

If we now formally take $h \to 0$ we arrive at the dynamical system

$$\begin{cases} \dfrac{dS_t}{dt} = -\alpha S_t I_t \\ \dfrac{dI_t}{dt} = \alpha S_t I_t - \rho I_t. \end{cases} \tag{1}$$

System (1) is the classic Kermack-McKendrick deterministic model. To investigate the infection spread under this model, we only need to consider nonnegative solutions for S, I, and R. The epidemic stops when $I_t = 0$ for the first time. Before we justify the approximation of the general epidemic process by the Kermack-McKendrick deterministic model, let us look at system (1) more closely.

Suppose $I_0 > 0$, $S_0 > 0$, and $R_0 = 0$ (this guarantees that $R_t \geq 0$ for all $t > 0$). The key questions is, given parameters α, ρ and the initial number of infecteds and susceptibles, whether the infection spreads and how it develops with time. Notice that S_t decreases with t, and

$$\frac{dI_t}{dt} = I_t(\alpha S_t - \rho) \begin{cases} \leq I_t(S_0 - \rho) \leq 0 \text{ for all } t > 0, & \text{if } S_0 \leq \rho/\alpha \\ > 0 \text{ for some } t > 0, & \text{if } S_0 > \rho/\alpha. \end{cases}$$

In the case when $S_0 \leq \rho/\alpha$, the number of infecteds monotonically decreases with time, that is no epidemic can occurs. By an epidemic we mean the situation when $I_t > I_0$ for some $t > 0$. On the other hand, when $S_0 > \rho/\alpha$, $dI_t/dt > 0$ at least initially, and the number of infecteds increases in the beginning. We observe the *threshold phenomena* at $S_0 = \rho/\alpha$, or qualitatively different infection spread above and below this level.

The critical parameter $R_0 \equiv \alpha S_0/\rho$ is called the *basic reproduction number*, and is defined as the number of secondary infections introduced by one primary infection into a wholly susceptible population. We will see that much like in the case of system (1) in many epidemic models $R = 1$ is the critical value; $R < 1$ implies no epidemic and $R > 1$ that an epidemic is possible.

1.2.1 Law of Large Numbers for General Epidemic Processes

We will now define and show rigorously why the trajectories of system (1) approximate general epidemic processes. This amounts to proving the Law of Large Numbers for the family of stochastic processes defined by the general epidemic processes indexed by the population size N. Let $s_t^N = S_t/N$, $i_t^N = I_t/N$, and $r_t^N = R_t/N = 1 - s_t^N - i_t^N$ be the proportion of susceptibles, infecteds, and recovered respectively. Let the infection rate α depend on N in such a way that $\alpha = \theta/N$ for some constant $\theta > 0$. In terms of population proportions system (1) becomes (dividing both equations by N)

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

Trajectories of the dynamical system (2) lie in the triangle

$$K \equiv \{(s,i) : s + i \leq 1, s \geq 0, i \geq 0\}.$$

The general epidemic population proportion processes $\gamma^N = (s_t^N, i_t^N)_{t \geq 0}$ take values in

$$K^N \equiv \{(s,i) : s + i \leq 1, (Ns, Ni) \in \mathbb{Z}_+^2\} \subset K.$$

Before we state the main theorem of this section, we describe a construction of the general epidemic processes as time-changed Poisson processes that will be used in the proof.

Let $Y(t)$ be a rate one Poisson process. Then $Y(ct)$ is a rate c Poisson process. Moreover, to construct a process with rate c_1 until time t_1 and rate c_2 thereafter, we could let

$$Z(t) = \begin{cases} Y(c_1 t), & t \leq t_1 \\ Y\big(c_1 t_1 + c_2(t - t_1)\big), & t > t_1 \end{cases} \quad = \quad Y\left(\int_0^t \lambda(u)du\right),$$

where

$$\lambda(u) = \begin{cases} c_1, & u \leq t_1 \\ c_2, & u > t_1 \end{cases}$$

is the rate function.

We can construct the general epidemic processes in a similar manner. Between transitions of type $S \to I$ and $I \to R$ the rates are constant and are equal to $N\theta s_u i_u$ and $N\rho i_u$ respectively. Let Y_1 and Y_2 be two independent rate one Poisson processes. Then for $N \in \mathbb{N}$

$$Y_1\left(\int_0^t N\theta s_u i_u du\right) \quad \text{and} \quad Y_2\left(\int_0^t N\rho i_u du\right)$$

count transitions of type $S \to I$ and $I \to R$ respectively. Let $y_i(t) = \frac{1}{N}Y_i(Nt)$, $i = 1, 2$. Then

$$\begin{aligned} s_t^N &= s_0^N - y_1\left(\int_0^t \theta s_u^N i_u^N du\right) \\ i_t^N &= i_0^N + y_1\left(\int_0^t \theta s_u^N i_u^N du\right) - y_2\left(\int_0^t \rho i_u^N du\right) \end{aligned} \tag{3}$$

is a version of a general epidemic process constructed on the same probability space for every population size $N \in \mathbb{N}$ (for rigorous details on this construction see Ethier and Kurtz (1986)).

We are now ready to state the main result of this section. Let $\gamma_t^N = (s_t^N, i_t^N)$ be the general epidemic processes constructed by (3), and let $\bar{\gamma}_t = (\bar{s}_t, \bar{i}_t)$ be the trajectories of (2).

Proposition 1. If $\lim_{N\to\infty} \gamma_0^N = \bar{\gamma}_0$ then for any $T > 0$

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

where $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^2$.

Proof. Let $\tilde{y}_i(s) = y_i(s) - s$ be the centered (and scaled) Poisson processes. Omitting index N to simplify the notation

$$|s_t - \bar{s}_t| \leq |s_0 - \bar{s}_0| + \left|\tilde{y}_1\left(\int_0^t \theta s_u i_u du\right)\right| + \left|\int_0^t \theta(s_u i_u - \bar{s}_u \bar{i}_u) du\right|,$$

and

$$|i_t - \bar{\imath}_t| \le |i_0 - \bar{\imath}_0| + \left|\tilde{y}_1\left(\int_0^t \theta s_u i_u du\right)\right| + \left|\tilde{y}_2\left(\int_0^t \rho i_u du\right)\right| + \left|\int_0^t \theta(s_u i_u - \bar{s}_u\bar{\imath}_u)du\right| + \left|\int_0^t \rho(i_u - \bar{\imath}_u)du\right|.$$

Recall that the strong LLN for Poisson processes says that for any $v > 0$

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

Let

$$\varepsilon(t) = \max\left\{\tilde{y}_1\left(\int_0^t \theta s_u i_i du\right), \tilde{y}_2\left(\int_0^t \rho i_i du\right)\right\}.$$

Since $\int_0^t \theta s_u i_u du \le \theta T$ and $\int_0^t \rho i_u du \le \rho T$ for all $t \le T$ the SLLN for Poisson processes implies that $\lim_{N\to\infty} \sup_{t\le T} \varepsilon(t) = 0$ a.s. $i = 1, 2$.

There exists an $M > 0$ such that

$$|\theta s_u i_u - \theta \bar{s}_u \bar{\imath}_u| \le M\|\gamma_u - \bar{\gamma}_u\| \quad \text{and} \quad |\rho i_u - \rho \imath_u| \le M\|\gamma_u - \bar{\gamma}_u\|$$

Combining the estimates we get

$$|s_t - \bar{s}_t| \le |s_0 - \bar{s}_0| + \varepsilon(t) + M\int_0^t \|\gamma_u - \bar{\gamma}_u\|du, \quad \text{and}$$
$$|i_t - \bar{\imath}_t| \le |i_0 - \bar{\imath}_0| + \varepsilon(t) + M\int_0^t \|\gamma_u - \bar{\gamma}_u\|du$$

so that

$$\|\gamma_t - \bar{\gamma}_t\| \le 2\left(\|\gamma_0 - \bar{\gamma}_0\| + \varepsilon(t) + M\int_0^t \|\gamma_u - \bar{\gamma}_u\|du\right).$$

Gronwall's inequality (see below) implies that

$$\|\gamma_t - \bar{\gamma}_t\| \le 2\big(\|\gamma_0 - \bar{\gamma}_0\| + \varepsilon(t)\big)e^{2MT} \to 0$$

as $N \to \infty$. □

Gronwall's Inequality

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

$$0 \leq f(t) \leq \epsilon + M \int_0^t f(s)ds, \qquad t \leq T$$

then $f(t) \leq \epsilon e^{Mt}$, $t \leq T$.

The result says that under some conditions for large population sizes N the population proportions of the general epidemic processes are well approximated by the trajectories of the deterministic system (2).

References

Bailey, N. T. J. (1975). *The mathematical theory of infectious diseases and its applications.* 2nd ed. Hafner Press, New York.

Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes. Characterization and Convergence*. Wiley, New York.

Kermack, W. O. and McKendrick, A. G. (1927). Contributions to the mathematical theory of epidemics, part i. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematics.* **115** 700–721.

Kermack, W. O. and McKendrick, A. G. (1932). Contributions to the mathematical theory of epidemics, ii - the problem of endemicity. *Proceedings of the Royal Society of Edinburgh. Section A. Mathematics.* **138** 55–83.

Kermack, W. O. and McKendrick, A. G. (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.