

## The SIR infection model

Assume that a new infectious disease spreads in a population. We are interested in the dynamics of infected individuals, people that have recovered from the disease, and potentially further groups (e.g. people who are infected, but do not yet spread the disease during an incubation period, or people in the hospital or the number of deceased). Models for such disease dynamics are designed and discussed in the field of mathematical ecology. The most basic models go back to a series of papers by William Kermack and Anderson Gray McKendrick (a student of Ronald Ross) from 1927 to 1933. Kermack-McKendrick theory is still the conceptual basis of more complex epidemiological models today. This holds, in particular, for the SIR infection model.

For the basic SIR model, we assume that individuals can get a disease only once and consider three groups of individuals in an otherwise unstructured population: susceptibles  $S(t)$  that can be infected by the disease, currently infected people  $I(t)$  that are also infectious for others, and a removed class  $R(t)$  of people who have had the disease and cannot be infected any more – either because they have recovered and are resistant or because they are dead. In the simplest version of the model, we assume that the disease dynamics unfolds rapidly, such that births and deaths in the population – other than deaths due to the disease itself – can be ignored. We then have a constant population size

$$S(t) + I(t) + R(t) = N = \text{const.}$$

The disease dynamics can be described by the following set of ODEs,

$$\dot{S} = -\frac{c}{N} SI \tag{1a}$$

$$\dot{I} = \frac{c}{N} SI - rI \tag{1b}$$

$$\dot{R} = rI, \tag{1c}$$

where  $c$  is the average number of “transmission-allowing” contacts per individual and per time unit and  $r$  is the recovery rate, or more strictly the rate at which previously infectious individuals become non-infectious (which may also be because they are completely quarantined or because they die). We have initial conditions

$$S(0) = S_0 \quad ; \quad I(0) = I_0 = N - S_0 \quad ; \quad R(0) = 0.$$

Although we have three dynamical variables, the constraint of a constant population size reduces the dimension of the dynamical system to two. Note also that the dynamics of  $S$  and  $I$  does not depend on  $R(t)$  at all. We thus only need to consider the first two ODEs in (1). Comparing these with (??) we see that they are just a special case of the Lotka-Volterra predator-prey system, but without any intrinsic growth of the prey (= the susceptible population). We further observe the following:

- The number of susceptibles is monotonically declining and confined to the interval  $[0; S_0]$ .

- The combined number of susceptibles and infected individuals,  $S(t) + I(t)$ , is also monotonically declining and confined to the interval  $[0; N]$ .
- Calculating isoclines, it is easy to see that the ODE system corresponds to a degenerate case: all points on the axis  $I = 0$  are equilibrium points and there are no equilibria with  $I > 0$ . We thus see that the infected class will always be empty in the long term,  $I(\infty) = 0$  (strictly, this follows because  $S(t) + I(t)$  is a Lyapunov function, see the section on global stability below).

There are three questions of applied relevance that are usually asked for the SIR dynamics, which we will discuss below.

**Will the epidemic break out?** Once the disease has entered the population, the most basic question is whether it is able to spread and the number of infected will grow under community transmission. We have

$$\dot{I}(t) = \left( \frac{c}{N} S(t) - r \right) I(t)$$

and thus  $\dot{I} > 0$  iff

$$\mathcal{R}(t) = \frac{cS(t)}{rN} > 1. \quad (2)$$

$\mathcal{R}(t)$  is called the *effective reproductive ratio*<sup>1</sup>. Here,  $cS(t)/N$  is the average number of newly infected individuals per infected individual per time interval (e.g. per day). The rate  $r$  in the denominator defines the relevant time scale, i.e.  $T_r = 1/r$  is the expected time that it takes for an infectious individual to enter the “removed” class, where it is no longer infectious.  $\mathcal{R}(t)$  is thus simply the average number of individuals that are newly infected by each infected individual. Obviously, this quantity depends on the current number  $S(t)$  of susceptible individuals. At the start of an outbreak, we usually have  $I_0 \ll N$  and thus  $S_0 \approx N$ . This motivates the definition of the *basic reproductive ratio* (also: *basic reproduction number*)

$$\mathcal{R}_0 = \frac{c}{r} = cT_r \quad [\approx \mathcal{R}(0) \text{ for } S_0 \approx N]. \quad (3)$$

Note that  $\mathcal{R}_0$  is just the average number of contacts during the time an individual is infectious. In contrast to  $\mathcal{R}(t)$  this is a fixed model parameter that does not change under the dynamics. A value of  $\mathcal{R}_0 > 1$  indicates that a disease can spread in a population without immunity. The initial exponential increase in the number of infected individuals then reads

$$I(t) = \exp[(\mathcal{R}_0 - 1)t/T_r] I_0$$

and  $\mathcal{R}_0$  can be estimated from the daily growth factor as

$$\mathcal{R}_0 = 1 + T_r \log \left[ \frac{I(t+1)}{I(t)} \right] \approx 1 + T_r \frac{I(t+1) - I(t)}{I(t)}.$$

---

<sup>1</sup>Unfortunately, both the “removed” class and the reproductive ratio are commonly denoted by ‘capital-R’. We use the calligraphic symbol for the reproductive ratio to distinguish both quantities.

**The maximal number of infected individuals** In the case of an epidemic ( $\mathcal{R}_0 > 1$ ), the number of infected individuals  $I(t)$  will first increase and the number of susceptibles  $S(t)$  decrease until the effective reproductive ratio reaches

$$\mathcal{R}(t) = \frac{cS(t)}{rN} = 1,$$

With further decline of  $S(t)$ , also  $I(t)$  declines and eventually reaches zero. For practical reasons (e.g., hospital beds required), one is often interested in the maximal number of infected individuals  $I_{\max}$ . We can express the change in  $I$  as a function of  $S$  as

$$\frac{dI}{dS} = \frac{dI/dt}{dS/dt} = \frac{(cS - rN)I}{-cSI} = \mathcal{R}_0 \frac{rN}{cS} - 1 \quad (4)$$

for  $I \neq 0$ . Integrating, and with the initial condition  $I(S_0) = I_0$ , we obtain

$$I(S) = \frac{r}{c}N \log[S] - S + I_0 + S_0 - \frac{r}{c}N \log[S_0] = N - S + \frac{r}{c}N \log[S/S_0]. \quad (5)$$

The maximum is obtained for  $S = Nr/c = N/\mathcal{R}_0$ , thus

$$\frac{I_{\max}}{N} = 1 - \frac{1 - \log[N/(S_0\mathcal{R}_0)]}{\mathcal{R}_0} \approx 1 - \frac{1 + \log[\mathcal{R}_0]}{\mathcal{R}_0}, \quad (6)$$

for  $S_0 \approx N$ . In this limit, the maximal proportion of infected individuals in the population is thus uniquely determined by  $\mathcal{R}_0$ . It ranges from 0 (for  $\mathcal{R}_0 = 1$ ) to 1 (for  $\mathcal{R}_0 \rightarrow \infty$ ). If  $\mathcal{R}_0$  is only slightly larger than 1, we obtain  $I_{\max}/N \approx (\mathcal{R}_0 - 1)^2/2$  (the maximal number of infected individuals decreases quadratically as  $\mathcal{R}_0$  approaches 1).

**How many individuals remain uninfected?** We know that ultimately the number of infected individuals must tend to zero,  $I_\infty = 0$ . However, since all points on the  $I = 0$  axis are fixed points it is not clear how many individuals (if any) remain uninfected at the end of the epidemic. This number  $S_\infty$  is of obvious applied interest, in particular if the disease is severe. From the first and the third equation in (1), we can write

$$\frac{dS}{dR} = \frac{-cS}{rN} = -\mathcal{R}_0 \frac{S}{N}$$

and thus

$$S(R) = S_0 \exp\left[-\mathcal{R}_0 \frac{R}{N}\right].$$

Since  $S_\infty = N - R_\infty$  we obtain  $S_\infty$  as the solution of the transcendent equation

$$S_\infty = S_0 \exp[-\mathcal{R}_0(1 - S_\infty/N)] > 0. \quad (7)$$

We thus see that there is always a fraction of uninfected individuals remaining at the end of the outbreak. With  $S_0 \approx N$  the proportion  $s_\infty := S_\infty/N$  of individuals that escape the epidemic is the smaller root of

$$s_\infty = \exp[\mathcal{R}_0(s_\infty - 1)].$$

For  $\mathcal{R}_0$  running from 1 to  $\infty$ ,  $s_\infty$  declines from 1 to zero. For  $\mathcal{R}_0$  close to 1, we can approximate the percentage  $r_\infty$  of individuals that are affected as

$$r_\infty = 1 - s_\infty \approx \frac{2(\mathcal{R}_0 - 1)}{\mathcal{R}_0^2} \approx 2(\mathcal{R}_0 - 1).$$