

# Solutions to: A primer on the use of probability generating functions in infectious disease modeling

Joel C. Miller

## 1 Introduction

**Exercise 1.1** Except for the Poisson distribution handled in Example 1.1, derive the PGFs shown in Table 1 directly from the definition  $f(x) = \sum_i r_i x^i$ .

For the negative binomial, it may be useful to use the binomial series:

$$(1 + \delta)^\eta = 1 + \eta\delta + \frac{\eta(\eta - 1)}{2!}\delta^2 + \dots + \frac{\eta(\eta - 1)\dots(\eta - i + 1)}{i!}\delta^i + \dots$$

using  $\eta = -\hat{r}$  and  $\delta = -px$ .

### Solution 1.1

a. Uniform:  $p(\lambda) = 1$ , so the PGF is simply  $0 + 0 + \dots + 1x^\lambda + 0 + 0 + \dots = x^\lambda$ .

b. Binomial: We have  $r_i = \binom{n}{i} p^i q^{n-i}$ . So

$$\begin{aligned}\sum_i r_i x^i &= \sum_i \binom{n}{i} p^i q^{n-i} x^i \\ &= (px + q)^n\end{aligned}$$

by the binomial theorem.

c. Geometric: We have  $r_i = q^i p$ . So

$$\begin{aligned}\sum_i r_i x^i &= \sum_i p q^i x^i \\ &= p \sum_i (qx)^i \\ &= \frac{p}{1 - qx}\end{aligned}$$

by the sum of a geometric series.

d. Negative binomial: We have  $r_i = \binom{i + \hat{r} - 1}{i} q^{\hat{r}} p^i$ . So

$$\begin{aligned}\sum_i r_i x^i &= \sum_i \binom{i + \hat{r} - 1}{i} q^{\hat{r}} p^i x^i \\ &= q^{\hat{r}} \sum_i \binom{i + \hat{r} - 1}{i} p^i x^i \\ &= q^{\hat{r}} (1 - px)^{-\hat{r}} \\ &= \left(\frac{q}{1 - px}\right)^{\hat{r}}\end{aligned}$$

**Exercise 1.2** Consider the binomial distribution with  $n$  trials, each having success probability  $p = \lambda/n$ . Using Table 1, show that the PGF for the binomial distribution converges to the PGF for the Poisson

distribution in the limit  $n \rightarrow \infty$ , if  $\lambda$  is fixed.

**Solution 1.2** We have  $(px + q)^n$  where  $p = \lambda/n$  and  $q = 1 - p$ . Taking  $n$  to be large gives

$$\begin{aligned}\lim_{n \rightarrow \infty} (q + px)^n &= \lim_{n \rightarrow \infty} \left(1 + \frac{\lambda(x-1)}{n}\right)^n \\ &= e^{\lambda(x-1)}\end{aligned}$$

## 2 Discrete-time spread of a simple disease: early time

**Exercise 2.1 Monotonicity of  $\alpha_g$**

- By considering the biological interpretation of  $\alpha_g$ , explain why the sequence of inequalities  $0 = \alpha_0 \leq \alpha_1 \leq \dots \leq 1$  should hold. That is, explain why  $\alpha_0 = 0$ , why the  $\alpha_i$  form a monotonically increasing sequence, and why all of them are at most 1.
- Show that  $\alpha_g$  therefore converges to some non-negative limit  $\alpha$  that is at most 1 and that  $\alpha = \mu(\alpha)$ .
- Use Property A.9 to show that if  $\mu(0) \neq 0$  there exists a unique  $\alpha < 1$  solving  $\alpha = \mu(\alpha)$  if and only if  $\mathcal{R}_0 = \mu'(1) > 1$ .
- Assuming  $\mu(0) \neq 0$ , use Property A.9 to show that if  $\mathcal{R}_0 > 1$  then  $\alpha_g$  converges to the unique  $\alpha < 1$  solving  $\alpha = \mu(\alpha)$ , and otherwise  $\alpha_g$  converges to 1.

**Solution 2.1**

- In generation 0 there is an infected individual. So the disease is not extinct. Thus  $\alpha_0 = 0$ . If the outbreak is extinct at generation  $g$ , it remains extinct at later generations. So  $\alpha_g \leq \alpha_{g-1}$ . The eventual extinction probability is at most 1.
- Any sequence of numbers that is increasing and bounded from above must have a limit that is at most that bound. So it converges to some limit  $\alpha$ . Since  $\mu(\alpha_g) = \alpha_{g+1}$ , we conclude that  $\mu(\alpha_g)$  must converge to the same limit. Thus since  $\mu$  is continuous,  $\mu(\alpha) = \alpha$ .
- The assumptions of Property A.9 hold, with  $\mathcal{R}_0$  playing the role of  $f'(1)$ . The property states that if  $\mathcal{R}_0 \leq 1$  the only solution is  $\alpha = 1$ , while if  $\mathcal{R}_0 > 1$  there is exactly one other solution in  $(0, 1)$ .
- The second part of Property A.9 states that for  $\mathcal{R}_0 > 1$ , the solution must converge to this unique  $\alpha < 1$ . If  $\mathcal{R}_0 \leq 1$ , the above observations show that it must converge to a solution to  $\alpha = \mu(\alpha)$ , and the only possible choice is  $\alpha = 1$ .

**Exercise 2.2** Use Theorem 2.2 to prove Theorem 2.1.

**Solution 2.2** We only need the first two parts of Theorem 2.2. We have  $\alpha = \lim_{g \rightarrow \infty} \alpha_g = \lim_{g \rightarrow \infty} \mu(\alpha_g)$ . Because  $\mu$  is continuous we have  $\lim_{g \rightarrow \infty} \mu(\alpha_g) = \mu(\lim_{g \rightarrow \infty} \alpha_g) = \mu(\alpha)$ . So  $\alpha = \mu(\alpha)$ .

**Exercise 2.3** Show that if  $\mu(0) = 0$ , then  $\lim_{g \rightarrow \infty} \alpha_g = 0$ . By referring to the biological interpretation of  $\mu(0) = 0$ , explain this result.

**Solution 2.3** Let  $\mu(0) = 0$  and consider the smallest  $g \geq 0$  such that  $\alpha_{g+1} \neq 0$  (if it exists). Then  $\alpha_{g+1} = \mu(\alpha_g) = \mu(0) = 0$ . So no such  $g$  exists.

Biologically,  $\mu(0) = 0$  means that every individual has at least one offspring. Thus if we start with one individual, we can never end up with zero.

**Exercise 2.4** Find all PGFs  $\mu(y)$  with  $\mathcal{R}_0 \leq 1$  and  $\mu(0) = 0$ . Why were these excluded from Theorem 2.2?

**Solution 2.4** If  $\mu(y) = \sum_{i=0}^{\infty} p_i x^i$  is a PGF and  $\mu(0) = 0$ , then  $p_0 = 0$  and  $\mu(y) = \sum_{i=1}^{\infty} p_i x^i$  so  $\mathcal{R}_0 = \mu'(1) = \sum_{i=1}^{\infty} i p_i$ . This is  $\geq \sum_{i=1}^{\infty} p_i = 1$ , with equality only if  $p_i = 0$  for  $i > 1$ . Thus the only such function is  $\mu(y) = y$ .

This corresponds to each individual having exactly one offspring. So starting with one infection, at each generation there will remain exactly one infection, and there is no chance for extinction. All other cases with  $\mathcal{R}_0 \leq 1$  have a nonzero chance of having zero offspring, and thus extinction is possible (and in fact inevitable).

**Exercise 2.5 Larger initial conditions**

Assume that disease is introduced with  $m$  infections rather than just 1, or that it is not observed by surveillance until  $m$  infections are present. Assume that the offspring distribution PGF is  $\mu(y)$ .

- a. If  $m$  is known, find the extinction probability.
- b. If  $m$  is unknown but its distribution has PGF  $h(y)$ , find the extinction probability.

**Solution 2.5** Let  $\alpha$  be the extinction probability from one individual.

- a. The extinction probability given  $m$  initial infections is  $\alpha^m$ .
- b. Let  $h(x) = \sum_m q_m x^m$ . The probability of extinction is  $\sum_m q_m \alpha^m = h(\alpha)$ .

**Exercise 2.6 Extinction probability**

Consider a disease in which  $p_0 = 0.1$ ,  $p_1 = 0.2$ ,  $p_2 = 0.65$ , and  $p_3 = 0.05$  with a single introduced infection.

- a. Numerically approximate the probability of extinction within 0, 1, 2, 3, 4, or 5 generations up to five significant digits (assuming an infinite population).
- b. Numerically approximate the probability of eventual extinction up to five significant digits (assuming an infinite population).
- c. A surveillance program is being introduced, and detection will lead to a response. But it will not be soon enough to affect the transmissions from generations 0 and 1. From then on  $p_0 = 0.3$ ,  $p_1 = 0.4$ ,  $p_2 = 0.3$ , and  $p_3 = 0$ . Numerically approximate the new probability of eventual extinction after an introduction in an unbounded population [be careful that you do the function composition in the right order – review Properties A.1 and A.8].

**Solution 2.6** Define  $\mu_1(y) = \sum_i p_i y^i = 0.1 + 0.2y + 0.65y^2 + 0.05y^3$ .

- a. The extinction probability after 0 generations is  $\alpha_0 = 0$  and for  $g > 0$ , the extinction probability after  $g$  generations is  $\alpha_g = \mu_1(\alpha_{g-1})$ . So by iteratively applying  $\mu_1$  to 0 we get

$$\begin{aligned} \alpha_0 &= 0 \\ \alpha_1 &= \mu_1(\alpha_0) = 0.1 \\ \alpha_2 &= \mu_1(\alpha_1) = 0.12655 \\ \alpha_3 &= \mu_1(\alpha_2) \approx 0.13582 \\ \alpha_4 &= \mu_1(\alpha_3) \approx 0.13928 \\ \alpha_5 &= \mu_1(\alpha_4) \approx 0.14060 \end{aligned}$$

- b. Repeatedly applying  $\mu_1$  to the result quickly yields convergence to  $\alpha \approx 0.14143$ .
- c. Let  $\mu_2(y) = 0.3 + 0.4y + 0.3y^2$ . Assuming that there are still infected individuals when the intervention is introduced. For each of them, the probability that all offspring die out is  $\lim_{g \rightarrow \infty} \mu_2^{[g]}(0) \approx 0.9688$ . The distribution of the number infected when the intervention is introduced is  $\mu_1(\mu_1(y))$ . So the probability of extinction is  $\mu_1(\mu_1(0.9688)) \approx 0.9185$ .

**Exercise 2.7** We look at two inductive derivations of  $\Phi_g(y) = \mu^{[g]}(y)$ . They are similar, but when adapted to the continuous-time dynamics we study later, they lead to two different models. We take as given that  $\Phi_{g-1}(y)$  gives the distribution of the number of infections caused after  $g-1$  generations starting from a single case. One argument is based on discussing the results of outcomes attributable to the infectious individuals of generation  $g-1$  in the next generation. The other is based on the outcomes indirectly attributable to the infectious individuals of generation 1 through their descendants after another  $g-1$  generations.

- a. Explain why Property A.8 shows that  $\Phi_g(y) = \Phi_{g-1}(\mu(y))$ .
- b. (without reference to a) Explain why Property A.8 shows that  $\Phi_g(y) = \mu(\Phi_{g-1}(y))$ .

**Solution 2.7** a. The PGF for the number infected in generation  $g-1$  is  $\Phi_{g-1}(y)$ . Each of the individuals who are infected in generation  $g-1$  make some sort of contribution (possibly 0) to generation  $g$ . The PGF for this contribution is  $\mu(y)$ . Thus by Property A.8, the PGF for the combination of steps from generation 0 to  $g-1$  and from  $g-1$  to  $g$  is given by  $\Phi_{g-1}(\mu(y))$ .

- b. The PGF for the number infected in generation 1 is  $\mu(y)$ . If we trace the contributions of these individuals to generation  $g$ , we find that the distribution is the same as for the contribution of a generation 0 individual to generation  $g-1$ . So by Property A.8, the PGF for the combination of steps from generation 0 to 1 and from 1 to  $g$  is  $\mu(\Phi_{g-1}(y))$ .

**Exercise 2.8** Use Theorem 2.3 to prove the first part of Theorem 2.2.

**Solution 2.8** The probability that there are no infections at generation  $g$  is given by  $\Phi_g(0)$  (The expansion eliminates all terms except the coefficient of  $y^0$ ).

Simply observing  $\Phi_g(0) = \mu^{[g]}(0)$  finishes the proof.

**Exercise 2.9** How does Corollary 2.1 change if we start with  $k$  infections?

**Solution 2.9** The first part is simply multiplied by  $k$ .

For  $\langle I \rangle_g$ , the numerator is multiplied by  $k$ . The denominator is the probability that the disease persists to generation  $g$ . This becomes  $1 - \alpha_g^k$ . This is because to be extinct, all  $k$  introductions must independently go extinct, which occurs with probability  $\alpha_g^k$ .

**Exercise 2.10** Assume the PGF of the offspring size distribution is  $\mu(y) = (1 + y + y^2)/3$ .

- a. What offspring size distribution yields this PGF?
- b. Find the PGF  $\Omega_g(z)$  for the number of completed infections at 0, 1, 2, 3, and 4 generations [it may be helpful to use a symbolic math program once  $g > 2$ ].
- c. Check that for these cases, once  $g > r$ , the coefficient of  $z^r$  does not change.

**Solution 2.10**

- a. Each individual causes 0, 1, or 2 transmissions with equal probability.
- b. (i) We have  $\Omega_0(z) = 1$  [there are no completed infections at generation 0].  
 (ii) We have  $\Omega_1(z) = z\mu(1) = z$  [which means that there is a single completed infection at generation 1].  
 (iii) We have

$$\begin{aligned}\Omega_2(z) &= z\mu(\Omega_1(z)) \\ &= z\mu(z) \\ &= \frac{z + z^2 + z^3}{3}\end{aligned}$$

(iv) We have

$$\begin{aligned}
\Omega_3(z) &= z\mu(\Omega_2(z)) \\
&= z \frac{1 + \Omega_2(z) + \Omega_2(z)^2}{3} \\
&= \frac{9z + 3z^2 + 4z^3 + 5z^4 + 3z^5 + 2z^6 + z^7}{27} \\
&= \frac{z}{3} + \frac{z^2}{9} + \dots
\end{aligned}$$

(v) and finally

$$\begin{aligned}
\Omega_4(z) &= z\mu(\Omega_3(z)) \\
&= z \frac{1 + \Omega_3(z) + \Omega_3(z)^2 + \Omega_3(z)^3}{3} \\
&= \frac{1}{2187} (729z + 243z^2 + 162z^3 + 162z^4 + 216z^5 + 105z^6 + 154z^7 \\
&\quad + 121z^8 + 79z^9 + 52z^{10} + 37z^{11} + 22z^{12} + 10z^{13} + 4z^{14} + z^{15}) \\
&= \frac{z}{3} + \frac{z^2}{9} + \dots
\end{aligned}$$

- c.
- For  $r = 0$ , we see that from  $\Omega_1(z)$  onwards, the coefficient of  $z^0$  is 0.
  - For  $r = 1$ , we see that from  $\Omega_2(z)$  onwards, the coefficient of  $z$  is  $1/3$
  - For  $r = 2$ , we see that from  $\Omega_3(z)$  onwards, the coefficient of  $z^2$  is  $1/9$ .

**Exercise 2.11** By setting  $y = 1$ , use Theorem 2.5 to prove Theorem 2.4.

**Solution 2.11** We expect that  $\Pi_g(1, z) = \Omega(z)$ . This can be checked by noting that for given  $r$ , the sum  $\sum_i \pi_{i,r}(g) 1^i z^r$  is  $z^r \sum_i \pi_{i,r}(g) = z^r \omega_r(g)$ .

So Theorem 2.5 states that  $\Pi_g(1, z) = z\mu(\Pi_{g-1}(1, z))$  which becomes  $\Omega_g(z) = z\mu(\Omega_{g-1}(z))$ .

**Exercise 2.12** Redo example 2.10 if  $\hat{r}$  is a real number, rather than an integer. It may be useful to use the  $\Gamma$ -function, which satisfies  $\Gamma(x+1) = x\Gamma(x)$  for any  $x$  and  $\Gamma(n+1) = n!$  for integer  $n$ .

**Solution 2.12** The key observation is that  $\hat{r}j(\hat{r}j+1)\cdots(\hat{r}j+j-2)/(j-1)!$  becomes  $\Gamma(\hat{r}j+j-1)/\Gamma(\hat{r}j-1)(j-1)!$ . Thus we replace  $\binom{\hat{r}j+j-2}{j-1}$  in the expression to yield

$$\frac{1}{j} \frac{\Gamma(\hat{r}j+j-2)}{\Gamma(\hat{r}j-1)(j-1)!} q^{\hat{r}j} p^{j-1}$$

**Exercise 2.13** Except for the negative binomial case done in example 2.10, derive the probabilities in Table 6.

- a. For the Poisson distribution, use Property A.2.
- b. For the Uniform distribution, use Property A.2.
- c. For the Binomial distribution, use the binomial theorem:  $(a+b)^c = \sum_{i=0}^c \binom{c}{i} a^i b^{c-i}$ .
- d. For the Geometric distribution, follow example 2.10 (noting that  $p$  and  $q$  interchange roles).

**Solution 2.13** For each we need to find the coefficient of  $y^{j-1}$  in  $[\mu(y)]^j$  and then divide it by  $j$ .

- a.  $[\mu(y)]^j = e^{j\lambda(y-1)}$  The coefficient of  $y^{j-1}$  is

$$\begin{aligned} \frac{1}{(j-1)!} \left( \frac{d}{dy} \right)^{j-1} e^{j\lambda(y-1)} \Big|_{y=0} &= \frac{1}{j-1} (j\lambda)^{j-1} e^{j\lambda(y-1)} \Big|_{y=0} \\ &= \frac{(j\lambda)^{j-1}}{(j-1)!} e^{-\lambda j} \end{aligned}$$

Taking  $1/j$  times this yields  $\frac{(\lambda j)^{j-1}}{j!} e^{-\lambda j}$ .

- b. Note that for this to make sense  $\lambda$  must be a non-negative integer. We have  $[\mu(y)]^j = y^{\lambda j}$ . The first case we consider is  $j = 1$ ,  $\lambda = 0$ . Then we have  $[\mu(y)]^j = 1$ . Taking zero derivatives, setting  $y = 0$ , and dividing by  $(j-1)! = 1$  and  $j = 1$  yields 1.

Now consider  $\lambda \geq 1$ . Then  $\lambda j > j - 1$ . So after taking  $j - 1$  derivatives, we still have a factor of  $y$ , which when we set  $y = 0$  yields 0.

Now consider  $\lambda = 0$ , but  $j > 1$ . We have  $[\mu(y)]^j = 1$ . Taking  $j - 1 \geq 1$  derivatives yields 0.

- c. We have  $[\mu(y)]^j = (q + py)^{jn}$ . By the binomial theorem, the coefficient of  $y^{j-1}$  is  $\binom{jn}{j-1} q^{jn-j+1} p^{j-1}$ . Taking  $1/j$  times this yields the result
- d. We have  $[\mu(y)]^j = (p/(1-xy))^j$ . Following the same steps as in the negative binomial distribution, but taking  $\hat{r} = 1$  and interchanging  $p$  and  $q$ , we end up with  $\frac{1}{j} \binom{2j-2}{j-1} p^j q^{j-1}$ .

**Exercise 2.14** To help model continuous-time epidemics, section 3 will use a modified version of  $\mu$ , which in some contexts will be written as  $\hat{\mu}(y, z)$ . To help motivate the use of two variables, we reconsider the discrete case. We think of a recovery as an infected individual disappearing and giving birth to a recovered individual and a collection of infected individuals. Look back at the discrete-time calculation of  $\Omega_g$  and  $\Pi_g$ . Define a two-variable version of  $\mu$  as  $\mu(y, z) = z \sum_i r_i y^i = z\mu(y)$ .

- What is the biological interpretation of  $\mu(y, z) = z\mu(y)$ ?
- Rewrite the recursive relations for  $\Omega_g$  using  $\mu(y, z)$  rather than  $\mu(y)$ .
- Rewrite the recursive relations for  $\Pi_g$  using  $\mu(y, z)$  rather than  $\mu(y)$ .

The choice to use  $\mu(y, z)$  versus  $\mu(y)$  is purely a matter of convenience.

### Solution 2.14

- After a generation, an individual contributes  $\mu(y)$  new infections and 1 new recovery. This is captured by  $\mu(y, z) = z\mu(y)$ .
- $\Omega_g(z) = \mu(\Omega_{g-1}(z), z)$ .
- $\Pi_g(y, z) = \mu(\Pi_{g-1}(y, z), z)$ .

**Exercise 2.15** Consider Example 2.11. Assume that a third outbreak is observed with 4 infections. Calculate the probability of  $\Theta_1$  and  $\Theta_2$  given the data starting

- with the assumption that  $P(\Theta_1) = P(\Theta_2) = 0.5$  and  $X$  consists of the three observations  $j = 7$ ,  $j = 8$ , and  $j = 4$ .
- with the assumption that  $P(\Theta_1) = 0.6546$  and  $P(\Theta_2) = 0.3454$  and  $X$  consists only of the single observation  $j = 4$ .
- Compare the results and explain why they should have the relation they do.

**Solution 2.15**

- a. There are now three observations  $j = 7$ ,  $j = 8$ , and  $j = 4$ . Adapting the result in the example, we have

$$f(\Theta) = \left( \sum_{j=7,8,4} \log((\hat{r}j + j - 2)!) - \log(j!) - \log((\hat{r}j - 1)!) + \hat{r}j \log q + (j - 1) \log p \right) + \log 0.5.$$

We find  $f(\Theta_1) \approx -11.3978$  and  $f(\Theta_2) \approx -12.0858$ . So  $\hat{f}(\Theta_1) = 0$  and  $\hat{f}(\Theta_2) = -0.688$ . Then  $g(\Theta_1) = 1$  and  $g(\Theta_2) = 0.5026$ . We finally have

$$P(\Theta_1|X) = 0.6655, \quad P(\Theta_2|X) = 0.3345$$

- b. The difference appears at the beginning:

$$f(\Theta_1) = (\log((\hat{r}4 + 4 - 2)!) - \log(4!) - \log((\hat{r}4 - 1)!) + \hat{r}4 \log q + (4 - 1) \log p) + \log 0.6546$$

and

$$f(\Theta_2) = (\log((\hat{r}4 + 4 - 2)!) - \log(4!) - \log((\hat{r}4 - 1)!) + \hat{r}4 \log q + (4 - 1) \log p) + \log 0.3454.$$

Plugging in for  $\hat{r}$ , we get  $f(\Theta_1) = -3.326$  and  $f(\Theta_2) = -4.014$ . Once we find  $\hat{f}$  we find that it takes the same value as in the previous part, and so the results follow.

- c. If we update our beliefs with all of our observations, we should get the same final outcome. This should depend on whether we do it all at once, or sequentially (or even what order we do it sequentially).

**Exercise 2.16** Assume that we know a priori that the offspring distribution for a disease has a negative binomial distribution with  $p = 0.02$ . Assume that our a priori knowledge of  $\hat{r}$  is that it is an integer uniformly distributed between 1 and 80 inclusive. Given observed outbreaks of sizes 1, 4, 5, 6, and 10:

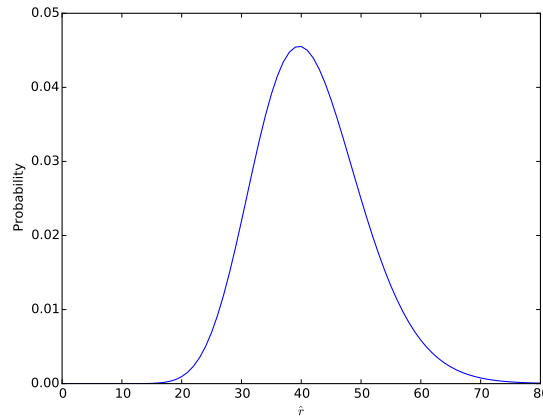
- a. For each  $\hat{r}$ , calculate  $P(\hat{r}|X)$  where  $X$  is the observed outbreak sizes. Plot the result.  
 b. Find the probability that  $\mathcal{R}_0 = \mu'(1)$  is greater than 1.

**Solution 2.16** The PGF is  $\left(\frac{0.98}{1-0.02x}\right)^{\hat{r}}$ . We start off with  $P(\Theta) = 1/80$  for all integers from 1 to 80. Given  $\hat{r}$  and  $p = 0.02$ , the probability of observing a particular size  $j$  is  $\frac{1}{j} \binom{\hat{r}+j-2}{j-1} 0.98^{\hat{r}j} 0.01^{j-1}$ .

- a. So  $P(X|\Theta)$  means the probability of observing sizes 1, 4, 5, 6, and 10 in five outbreaks given the value  $\hat{r}$ . For simplicity, we note that  $1 \cdot 4 \cdot 5 \cdot 6 \cdot 10 = 1200$  and  $1 + 4 + 5 + 6 + 10 = 26$  This is

$$P(1, 4, 5, 6, 10|\hat{r}) = \frac{1}{1200} 0.98^{26\hat{r}} 0.02^{26-5} \binom{\hat{r}-1}{0} \binom{4\hat{r}+2}{3} \binom{5\hat{r}+3}{4} \binom{6\hat{r}+4}{5} \binom{10\hat{r}+8}{9}$$

Following the steps we get



- b. By summing the probabilities over all  $\hat{r}$  for which  $\mathcal{R}_0 > 1$ , we find the probability that  $\mathcal{R}_0 > 1$  is 0.181

### 3 Continuous-time spread of a simple disease

#### Exercise 3.1 Extinction Probability

Let  $\beta$  and  $\gamma$  be given with  $\hat{\mu}(y) = (\beta y^2 + \gamma)/(\beta + \gamma)$ .

- Analytically find solutions to  $y = \hat{\mu}(y)$ .
- Assume  $\beta < \gamma$ . Find all solutions in  $[0, 1]$ .
- Assume  $\beta > \gamma$ . Find all solutions in  $[0, 1]$ .

#### Solution 3.1

- a.  $y(\beta + \gamma) = \beta y^2 + \gamma$  so

$$\beta y^2 - (\beta + \gamma)y + \gamma = 0$$

By the quadratic formula

$$y = \frac{\beta + \gamma \pm \sqrt{(\beta + \gamma)^2 - 4\beta\gamma}}{2\beta}$$

The radical becomes  $\sqrt{\beta^2 - 2\beta\gamma + \gamma^2} = (\beta - \gamma)$ . So

$$y = \frac{\beta + \gamma \pm (\beta - \gamma)}{2\beta} = \{1, \gamma/\beta\}$$

- If  $\beta < \gamma$ , then only 1 is in the interval.
- If  $\gamma > \beta$  then both are in the interval.

#### Exercise 3.2 Consistency with discrete-time formulation.

Although we have argued that a transmission in the continuous-time disease transmission case can be treated as if a single infected individual has two infected offspring and then disappears, this is not what actually happens. In this exercise we look at the true offspring distribution of an infected individual before recovery, and we show that the ultimate predictions of the two versions are equivalent.

Consider a disease in which individuals transmit at rate  $\beta$  and recover at rate  $\gamma$ . Let  $p_i$  be the probability an infected individual will cause exactly  $i$  new infections before recovering.

- Explain why  $p_0 = \gamma/(\beta + \gamma)$ .
- Explain why  $p_i = \beta^i \gamma / (\beta + \gamma)^{i+1}$ . So  $p_i$  form a geometric distribution.



- c. Show that  $\mu(y) = \sum_i p_i y^i$  can be expressed as  $\mu(y) = \gamma/(\beta + \gamma - \beta y)$ . [This definition of  $\mu$  without the hat corresponds to the discrete-time definition]
- d. Show that the solutions to  $y = \mu(y)$  are the same as the solutions to  $y = \hat{\mu}(y) = (\beta y^2 + \gamma)/(\beta + \gamma)$ . So the extinction probability can be calculated either way. (You do not have to find the solutions to do this, you can simply show that the two equations are equivalent).

**Solution 3.2**

- a. The event is either a recovery or a transmission, and the rates are  $\gamma$  and  $\beta$ . So the probability of recovery is the rate of recovery divided by the total rate.
- b. While the individual is susceptible, the probability that the next event is a transmission is  $\beta/(\beta + \gamma)$ . So the probability that the first  $i$  events are transmissions is  $\beta^i/(\beta + \gamma)^i$ . The probability that the next event is a recovery is  $\gamma/(\beta + \gamma)$ . The product of these is the probability that the first  $i$  events are transmissions and the next event is a recovery.
- c. Let  $r = \beta/(\beta + \gamma)$ . Then  $p_i = (1 - r)r^i$ . Then  $\sum_i p_i y^i$  is a geometric series:  $(1 - r) \sum_i (ry)^i = (1 - r)/(1 - ry)$ . Multiplying by  $1 = (\beta + \gamma)/(\beta + \gamma)$  gives the result.
- d. The solutions to  $y = \mu(y)$  are the solutions to  $y(\beta + \gamma - \beta y) = \gamma$ . The solutions to  $y = \hat{\mu}(y) = (\beta y^2 + \gamma)/(\beta + \gamma)$  are the solutions to  $y(\beta + \gamma) = \beta y^2 + \gamma$ . By moving  $\beta y^2$  to the other side of either equation we see that the equations are equivalent.

**Exercise 3.3 Relation with  $\mathcal{R}_0$**

Take  $\mu(y) = \gamma/(\beta + \gamma - \beta y)$  as given in exercise 3.2 and  $\hat{\mu} = (\beta y^2 + \gamma)/(\beta + \gamma)$ .

- a. Show that  $\mu'(1) \neq \hat{\mu}'(1)$  in general.
- b. Show that when  $\mathcal{R}_0 = \mu'(1) = 1$ , then  $\hat{\mu}'(1) = 1$ . So both are still threshold parameters.

**Solution 3.3**

- a.  $\mu'(y) = \gamma\beta/(\beta + \gamma - \beta y)^2$ , so  $\mu'(1) = \gamma\beta/\gamma^2 = \beta/\gamma$ . Similarly  $\hat{\mu}'(y) = 2\gamma y/(\beta + \gamma)$ , so  $\hat{\mu}'(1) = 2\gamma/(\beta + \gamma)$ . These are generally unequal.
- b. When  $\mathcal{R}_0 = 1$ , we have  $\beta = \gamma$ . So  $\mu'(1) = \gamma/\gamma = 1$  and  $\hat{\mu}'(1) = 2\gamma/(\gamma + \gamma) = 1$ . So the two are equal.

**Exercise 3.4 Revisiting eventual extinction probability.**

We revisit the results of exercise 3.1 using Eq. (17) (without solving it).

- a. By substituting for  $\hat{\mu}(\alpha)$ , show that  $\dot{\alpha} = (1 - \alpha)(\gamma - \beta\alpha)$ .

We have  $\alpha(0) = 0$ . Taking this initial condition and expression for  $\dot{\alpha}$ , show that

- b.  $\alpha \rightarrow 1$  as  $t \rightarrow \infty$  if  $\beta < \gamma$  (i.e.,  $\mathcal{R}_0 < 1$ ) and
- c.  $\alpha \rightarrow \gamma/\beta$  as  $t \rightarrow \infty$  if  $\beta > \gamma$  (i.e.,  $\mathcal{R}_0 > 1$ ).
- d. Set up (but do not solve) a partial fraction integration that would give  $\alpha(t)$  analytically.

**Solution 3.4**

- a. We have

$$\begin{aligned} \dot{\alpha} &= (\beta + \gamma) \left( \frac{\beta\alpha^2 + \gamma}{\beta + \gamma} - \alpha \right) \\ &= \beta\alpha^2 + \gamma - \alpha(\beta + \gamma) \\ &= (1 - \alpha)(\gamma - \beta\alpha) \end{aligned}$$

Starting from  $\alpha(0) = 0$ , we see that  $\dot{\alpha}(0) = \gamma > 0$ . So  $\alpha$  will increase and approach the smallest positive

equilibrium value. The two equilibria are at  $\alpha = 1$  and  $\alpha = \gamma/\beta$ .

- b. If  $\beta < \gamma$ , then the smaller equilibrium is  $\alpha = 1$ .
- c. If  $\beta > \gamma$ , then the smaller equilibrium is  $\alpha = \gamma/\beta$ .
- d. We have

$$\frac{\frac{d\alpha}{dt}}{(1-\alpha)(\gamma-\beta\alpha)} = 1$$

So using separation of variables, we have

$$\int \frac{1}{(1-\alpha)(\gamma-\beta\alpha)} d\alpha = \int 1 dt$$

Using partial fractions this becomes

$$\int \frac{A}{1-\alpha} + \frac{B}{\gamma-\beta\alpha} d\alpha = \int 1 dt$$

where  $A$  and  $B$  are chosen such that the sum in the integrand has a numerator of 1. Once  $A$  and  $B$  are found, both sides can be integrated analytically.

**Exercise 3.5** This exercise is intended to help with understanding the backward Kolmogorov equations.

Let  $\phi_i(t)$  denote the probability of having  $i$  active infections at time  $t$  given that at time 0 there was a single infection [ $\phi_1(0) = 1$ ]. We have  $\phi_0(t) = \alpha(t)$ . We extend the derivation of Eq. (16) to  $\phi_1$ . Assume  $\phi_0(t_0)$  and  $\phi_1(t_0)$  are known.

- a. Following the derivation of Eq. (16), approximate  $\phi_0(\Delta t)$ ,  $\phi_1(\Delta t)$ , and  $\phi_2(\Delta t)$  for small  $\Delta t$ .
- b. From biological grounds explain why if there are 0 infections at time  $\Delta t$  then there are also 0 infections at time  $t_0 + \Delta t$ .
- c. If there is 1 infection at time  $\Delta t$ , what is the probability of 1 infection at time  $t_0 + \Delta t$ ?
- d. If there are 2 infections at time  $\Delta t$ , what is the probability of 1 infection at time  $t_0 + \Delta t$ ?
- e. Write  $\phi_1(t_0 + \Delta t)$  in terms of  $\phi_0(t_0)$ ,  $\phi_1(t_0)$ ,  $\phi_1(\Delta t)$ , and  $\phi_2(\Delta t)$ .
- f. Using the definition of the derivative, find an expression for  $\dot{\phi}_1$  in terms of  $\phi_1(t)$  and  $\phi_2(t)$ .

**Solution 3.5**

- a. We have

$$\begin{aligned}\phi_0(\Delta t) &= \gamma\Delta t + o(\Delta t) \\ \phi_1(\Delta t) &= 1 - (\beta + \gamma)\Delta t + o(\Delta t) \\ \phi_2(\Delta t) &= \beta\Delta t + o(\Delta t)\end{aligned}$$

- b. If the disease is extinct at time  $\Delta t$ , then it remains extinct.
- c. This equals the probability of having 1 infection at time  $t_0$  if there is one infection at time 0. So it is  $\phi_1(t_0)$ .
- d. One of the infections has to have its descendants go extinct within  $t_0$  units of time, and the other must have 1 descendant after  $t_0$  units of time. So this is  $\phi_0(t_0)\phi_1(t_0) + \phi_1(t_0)\phi_0(t_0) = 2\phi_0(t_0)\phi_1(t_0)$
- e. So [using the additional fact that the probability of 3 or more infections at time  $\Delta t$  is  $o(\Delta t)$ ]

$$\phi_1(t_0 + \Delta t) = \phi_1(\Delta t)\phi_1(t_0) + 2\phi_2(\Delta t)\phi_0(t_0)\phi_1(t_0) + o(\Delta t)$$

f. So

$$\begin{aligned}
\dot{\phi}_1(t) &= \lim_{\Delta t \rightarrow 0} \frac{\phi_1(t + \Delta t) - \phi_1(t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{\phi_1(\Delta t)\phi_1(t) + 2\phi_2(\Delta t)\phi_0(t)\phi_1(t) + \mathcal{O}(\Delta t) - \phi_1(t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{[1 - (\beta + \gamma)\Delta t]\phi_1(t) + 2\beta\Delta t\phi_0(t)\phi_1(t) + \mathcal{O}(\Delta t) - \phi_1(t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{-(\beta + \gamma)\Delta t\phi_1(t) + 2\beta\Delta t\phi_0(t)\phi_1(t) + \mathcal{O}(\Delta t)}{\Delta t} \\
&= -(\beta + \gamma)\phi_1(t) + 2\beta\phi_0(t)\phi_1(t)
\end{aligned}$$

**Exercise 3.6** In this exercise we derive the PGF version of the forward Kolmogorov equations by directly calculating the rate of change of the probabilities of the states. Define  $\phi_j(t)$  to be the probability that there are  $j$  active infections at time  $t$ .

We have the forward Kolmogorov equations:

$$\dot{\phi}_j = \beta(j-1)\phi_{j-1} + \gamma(j+1)\phi_{j+1} - (\beta + \gamma)j\phi_j.$$

- Explain each term on the right hand side of the equation for  $\dot{\phi}_j$ .
- By expanding  $\dot{\Phi}(y, t) = \frac{\partial}{\partial t} \sum_j \phi_j y^j$ , arrive at Equation (18).

### Solution 3.6

- If there are  $j$  infected individuals, then a new transmission will result in  $j+1$  total infections and occurs at total rate  $\beta j$ . This appears as a loss to  $\phi_j$  at rate  $j\beta\phi_j$ , but a gain at rate  $(j-1)\beta\phi_{j-1}$  from the smaller state. Similarly a recovery will result in  $j-1$  total infections and occurs at total rate  $\gamma j$ . No other possibilities are considered in our model. This appears as a loss at rate  $j\gamma\phi_j$  and a gain at rate  $(j+1)\gamma\phi_{j+1}$ .
- We convert this into an equation for the PGF:

$$\begin{aligned}
\frac{\partial}{\partial t} \Phi(y, t) &= \frac{\partial}{\partial t} \sum_j \phi_j(t) y^j \\
&= \sum_j \dot{\phi}_j(t) y^j \\
&= \sum_j \beta(j-1)\phi_{j-1} y^j + \gamma(j+1)\phi_{j+1} y^j - (\beta + \gamma)j\phi_j y^j \\
&= \beta \sum_j (y^2 - y) \frac{\partial}{\partial y} \phi_{j-1} y^{j-1} + \gamma \sum_j \frac{\partial}{\partial y} \phi_{j+1} y^{j+1} - (\beta + \gamma) \sum_j y \frac{\partial}{\partial y} \phi_j y^j \\
&= [\beta(y^2 - y) + \gamma - (\beta + \gamma)y] \frac{\partial}{\partial y} \Phi(y, t) \\
&= (\beta + \gamma) [\hat{\mu}(y) - y] \frac{\partial}{\partial y} \Phi(y, t)
\end{aligned}$$

**Exercise 3.7** In this exercise we follow [3, 6] and derive the PGF version of the backward Kolmogorov equations by directly calculating the rate of change of the probabilities of the states. Define  $\phi_{ki}(t)$  to be the probability of  $i$  infections at time  $t$  given that there were  $k$  infections at time 0. Although we assume that at time 0 there is a single infection, we will need to derive the equations for arbitrary  $k$ .

- Explain why

$$\phi_{ki}(t + \Delta t) = \phi_{ki}(t) - k(\beta + \gamma)\phi_{ki}(t)\Delta t + k(\beta\phi_{(k+1)i}(t) + \gamma\phi_{(k-1)i}(t)) + \mathcal{O}(\Delta t)$$

for small  $\Delta t$ .

b. By using the definition of the derivative  $\dot{\phi}_{ki} = \lim_{\Delta t \rightarrow 0} \frac{\phi_{ki}(t+\Delta t) - \phi_{ki}(t)}{\Delta t}$ , find  $\dot{\phi}_{ki}$

Define  $\Phi(y, t|k) = \sum_i \phi_{ki} y^i$  to be the PGF for the number of active infections assuming that there are  $k$  initial infections.

c. Show that

$$\dot{\Phi}(y, t|1) = -(\beta + \gamma)\Phi(y, t|1) + \beta\Phi(y, t|2) + \gamma\Phi(y, t|0)$$

d. Explain why  $\Phi(y, t|k) = \Phi(y, t|1)^k$ .

e. Complete the derivation of Equation (19).

### Solution 3.7

a. If  $\Delta t$  is very small then to leading order either 0 or 1 event occurs in the first  $\Delta t$  units of time

- The probability of 0 events if we start with  $k$  infections is  $1 - k(\beta + \gamma)\Delta t + o(\Delta t)$ . In this case the probability of  $i$  infections at time  $t + \Delta t$  is the same as the probability of  $i$  infections at time  $t$  given the initial condition.
- The probability of 1 event occurring and it being an infection is  $k\beta\Delta t + o(\Delta t)$ . In this the probability of  $i$  infections at time  $t + \Delta t$  is the same as the probability of  $i$  infections at time  $t$  if we start with  $k + 1$  infections,  $\phi_{(k+1)i}$ .
- The probability of 1 event occurring and it being a recovery is  $k\gamma\Delta t + o(\Delta t)$ . Following the previous case, this corresponds to starting with  $k - 1$  infections,  $\phi_{(k-1)i}$ .

Adding these together gives the result.

b.

$$\dot{\phi}_{ki} = -k(\beta + \gamma)\phi_{ki} + k\beta\phi_{(k+1)i} + k\gamma\phi_{(k-1)i}$$

c.

$$\begin{aligned} \dot{\Phi}(y, t|1) &= \sum_i \dot{\phi}_{1i} y^i \\ &= \sum_i -(\beta + \gamma)\phi_{1i} y^i + \beta\phi_{2i} y^i + \gamma\phi_{0i} y^i \\ &= -(\beta + \gamma)\Phi(y, t|1) + \beta\Phi(y, t|2) + \gamma\Phi(y, t|0) \end{aligned}$$

d. This follows from  $k$  applications of Property A.6.

e. So we use  $\Phi(y, t|2) = \Phi(y, t|1)^2$  and  $\Phi(y, t|0) = 1$  to give

$$\begin{aligned} \dot{\Phi}(y, t|1) &= -(\beta + \gamma)\Phi(y, t|1) + \beta\Phi(y, t|1)^2 + \gamma \\ &= (\beta + \gamma) \left[ \frac{\beta\Phi(y, t|1)^2 + \gamma}{\beta + \gamma} - \Phi(y, t|1) \right] \\ &= (\beta + \gamma)[\mu(\hat{\Phi}(y, t|1)) - \Phi(y, t|1)] \end{aligned}$$

Replacing  $\Phi(y, t|1)$  with  $\Phi(y, t)$  completes the derivation.

**Exercise 3.8** Define  $\Phi(y, t|k)$  to be the PGF for the probability of having  $i$  infections at time  $t$  given  $k$  infections at time 0.

a. Explain why  $\Phi(y, t|k) = [\Phi(y, t)]^k$ .

- b. Show that if we substitute  $\Phi(y, t|k) = [\Phi(y, t)]^k$  in place of  $\Phi(y, t)$  in Eq. (18) the equation remains true with the initial condition  $y^k$ .
- c. Show that if we substitute  $\Phi(y, t|k) = [\Phi(y, t)]^k$  in place of  $\Phi(y, t)$  in equation (19) we do not get a true equation.

So Eq. (18) applies regardless of the initial condition, but Eq. (19) is only true for the specific initial condition of one infection.

### Solution 3.8

- a. This follows from  $k$  applications of Property A.6.
- b. On the left hand side, the substitution yields

$$\frac{\partial}{\partial t} \Phi(y, t)^k = k\Phi(y, t)^{k-1} \frac{\partial}{\partial t} \Phi(y, t)$$

and on the right hand side we get

$$(\beta + \gamma)[\hat{\mu}(y) - y] \frac{\partial}{\partial y} \Phi(y, t)^k = k\Phi(y, t)^{k-1} (\beta + \gamma)[\hat{\mu}(y) - y] \frac{\partial}{\partial y} \Phi(y, t)$$

The  $k\Phi(y, t)^{k-1}$  term on each side cancel, and we have a correct equation.

- c. On the left hand side the substitution yields

$$\frac{\partial}{\partial t} \Phi(y, t)^k = k\Phi(y, t)^{k-1} \frac{\partial}{\partial t} \Phi(y, t)$$

as before, but on the right hand side we get

$$(\beta + \gamma)[\hat{\mu}(\Phi(y, t)^k) - \Phi(y, t)^k] = \beta\Phi(y, t)^{2k} + \gamma - (\beta + \gamma)\Phi(y, t)^k$$

There is no common term we can cancel to get back to the original equation.

**Exercise 3.9** Let  $\Phi(y, t|k)$  be the PGF for the number of infections assuming there are initially  $k$  infections. Derive the backward Kolmogorov equation for  $\Phi(y, t|k)$ . Note that some of the  $\Phi$ s in the derivation above would correspond to  $\Phi(y, t|1)$  and some of them to  $\Phi(y, t|k)$ .

**Solution 3.9** We have

$$\begin{aligned} \dot{\Phi}(y, t|k) &= \sum_i \dot{\phi}_{ki} y^i \\ &= \sum_i [-k(\beta + \gamma)\phi_{ki} + k\beta\phi_{(k+1)i} + k\gamma\phi_{(k-1)i}] y^i \\ &= -k(\beta + \gamma)\Phi(y, t|k) + k\beta\Phi(y, t|k+1) + k\gamma\Phi(y, t|k-1) \end{aligned}$$

### Exercise 3.10 Comparison of the formulations

- a. Using Eq. (18) derive an equation for  $\dot{\alpha}$  where  $\alpha(t) = \Phi(0, t)$ . What, if any, additional information would you need to solve this numerically?
- b. Using Eq. (19), derive Equation (17) for  $\dot{\alpha}$  where  $\alpha(t) = \Phi(0, t)$ . What, if any, additional information would you need to solve this numerically?

### Solution 3.10

a. We know that  $\alpha(0) = 0$ . Substituting into Equation (18)

$$\begin{aligned}\dot{\alpha}(t) &= (\beta + \gamma)[\hat{\mu}(0) - 0] \frac{\partial}{\partial y} \Phi(0, t) \\ &= \gamma \frac{\partial}{\partial y} \Phi(0, t) \\ &= \gamma \phi_1(t)\end{aligned}$$

We would also need to know the probability of having one infection at any given time.

b. Again, we know that  $\alpha(0) = 0$ . Substituting into Equation (19) yields

$$\dot{\alpha} = (\beta + \gamma)[\hat{\mu}(\alpha) - \alpha]$$

We would not need any additional information to solve this

### Exercise 3.11 Full solution

a. Show that Eq. (19) can be written

$$\frac{\partial}{\partial t} \Phi(y, t) = (\gamma - \beta \Phi(y, t))(1 - \Phi(y, t))$$

b. Using partial fractions, set up an integral which you could use to solve for  $\Phi(y, t)$  analytically (you do not need to do all the algebra to solve it).

### Solution 3.11

a. If we substitute for  $\hat{\mu}$ , we get

$$\begin{aligned}\frac{\partial}{\partial t} \Phi(y, t) &= (\beta + \gamma) \left( \frac{\beta[\Phi(y, t)]^2 + \gamma}{\beta + \gamma} - \Phi(y, t) \right) \\ &= \beta[\Phi(y, t)]^2 + \gamma - (\beta + \gamma)\Phi(y, t)\end{aligned}$$

and

$$\begin{aligned}(\gamma - \beta \Phi(y, t))(1 - \Phi(y, t)) &= \gamma - \gamma \Phi(y, t) - \beta \Phi(y, t) + \beta[\Phi(y, t)]^2 \\ &= \beta[\Phi(y, t)]^2 - (\beta + \gamma)\Phi(y, t) + \gamma\end{aligned}$$

So these are equal.

b. We have

$$\int \frac{1}{(\gamma - \beta \Phi(y, t))(1 - \Phi(y, t))} d\Phi(y, t) = \int dt$$

We write the first integrand as  $\frac{1}{(\gamma - \beta \Phi(y, t))(1 - \Phi(y, t))} = \frac{A}{\gamma - \beta \Phi(y, t)} + \frac{B}{1 - \Phi(y, t)}$  and must solve for  $A$  and  $B$ . Then this is integrable.

**Exercise 3.12** Argue from their definitions that  $\Phi(y, t) = \Pi(y, z, t)|_{z=1}$ .

**Solution 3.12** We have  $\Phi(y, t) = \sum_i \phi_i(t) y^i$  where  $\phi_i(t)$  is the probability of  $i$  infections at time  $t$ . Simi-

larly

$$\begin{aligned}\Pi(y, z, t)|_{z=1} &= \sum_{i,r} \pi_{i,r} y^i z^r \Big|_{z=1} \\ &= \sum_i \sum_r \pi_{i,r} y^i 1^r \\ &= \sum_i \left( y^i \sum_r \pi_{i,r} \right)\end{aligned}$$

and because  $\sum_r \pi_{i,r} = \phi_i$ , the result follows.

**Exercise 3.13** Derive Theorem 3.3 from Theorem 3.4.

**Solution 3.13** Note that  $\Phi(y, t) = \Pi(y, 1, t)$ . Setting  $z = 1$  into the equation for  $\Pi$  yields the equation for  $\Phi$ .

**Exercise 3.14** Derive Theorem 3.5 from Theorem 3.4.

**Solution 3.14** Note that for  $z < 1$ ,  $\Omega_\infty(z) = \lim_{t \rightarrow \infty} \Pi(0, z, t)$  and for  $z = 1$  it is 1.

As  $t \rightarrow \infty$ , we must have  $\frac{\partial}{\partial t} P_i(0, z, t) = 0$  since the system approaches a disease-free state and thus all coefficients converge to a constant. For this to hold, Equation (22) yields

$$\lim_{t \rightarrow \infty} \hat{\mu}(\Pi(0, z, t), z) = \lim_{t \rightarrow \infty} \Pi(0, z, t)$$

Substituting with  $\Omega_\infty$  completes the result.

**Exercise 3.15** *Equivalence of continuous and discrete final size distributions.*

Show by direct substitution that if  $\Omega_\infty(z) = \hat{\mu}(\Omega_\infty(z), z)$  then  $\Omega_\infty(z) = z\mu(\Omega_\infty(z))$  where  $\mu(y) = \gamma/(\beta + \gamma - \beta y)$  is the PGF for the offspring distribution found in Exercise 3.2.

**Solution 3.15** We take

$$\begin{aligned}\Omega_\infty(z) &= \hat{\mu}(\Omega_\infty(z), z) \\ &= \frac{\beta[\Omega_\infty(z)]^2 + \gamma z}{\beta + \gamma}\end{aligned}$$

So

$$\begin{aligned}(\beta + \gamma)\Omega_\infty(z) &= \beta[\Omega_\infty(z)]^2 + \gamma z \\ \Omega_\infty(z)[\beta + \gamma - \beta\Omega_\infty(z)] &= \gamma z \\ \Omega_\infty(z) &= \frac{\gamma z}{\beta + \gamma - \beta\Omega_\infty(z)}\end{aligned}$$

**Exercise 3.16** We revisit the derivations of the usual mass action SIR ODEs. Following Example 3.3,

- Derive  $[\dot{S}]$  in terms of  $[SI]$ .
- Derive  $[\dot{I}]$  in terms of  $[SI]$  and  $[I]$ .
- Using  $[S] + [I] + [R] = N$ , derive  $[\dot{R}]$ .

**Solution 3.16**

a. We have

$$\begin{aligned}
[\dot{S}] &= \frac{\partial}{\partial t} \frac{\partial}{\partial x} \Xi(x, y, t) \Big|_{x=y=1} \\
&= \frac{\partial}{\partial x} \frac{\partial}{\partial t} \Xi(x, y, t) \Big|_{x=y=1} \\
&= \frac{\partial}{\partial x} \left( \frac{\beta(y^2 - xy)}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(x, y, t) + \gamma(1 - y) \frac{\partial}{\partial y} \Xi(x, y, t) \right) \Big|_{x=y=1} \\
&= -\frac{\beta y}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(x, y, t) + \frac{\beta(y^2 - xy)}{N} \frac{\partial^2}{\partial x^2} \frac{\partial}{\partial y} \Xi(x, y, t) + \gamma(1 - y) \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(x, y, t) \Big|_{x=y=1} \\
&= -\frac{\beta}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(1, 1, t) \\
&= -\frac{\beta}{N} [SI]
\end{aligned}$$

b. We have

$$\begin{aligned}
[\dot{I}] &= \frac{\partial}{\partial t} \frac{\partial}{\partial y} \Xi(x, y, t) \Big|_{x=y=1} \\
&= \frac{\partial}{\partial y} \left( \frac{\beta(y^2 - xy)}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(x, y, t) + \gamma(1 - y) \frac{\partial}{\partial y} \Xi(x, y, t) \right) \Big|_{x=y=1} \\
&= \frac{\beta(2y - x)}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(x, y, t) + \frac{\beta(y^2 - xy)}{N} \frac{\partial}{\partial x} \frac{\partial^2}{\partial y^2} \Xi(x, y, t) - \gamma \frac{\partial}{\partial y} \Xi(x, y, t) + \gamma(1 - y) \frac{\partial^2}{\partial y^2} \Xi(x, y, t) \Big|_{x=y=1} \\
&= \frac{\beta}{N} \frac{\partial}{\partial x} \frac{\partial}{\partial y} \Xi(1, 1, t) - \gamma \frac{\partial}{\partial y} \Xi(1, 1, t) \\
&= \frac{\beta}{N} [SI] - \gamma [I]
\end{aligned}$$

c. As  $[\dot{S}] + [\dot{I}] + [\dot{R}] = 0$ , we conclude

$$[\dot{R}] = -[\dot{S}] - [\dot{I}] = \gamma [I]$$

## 4 Large-time dynamics

### Exercise 4.1 Ancestor distribution for homogeneous well-mixed population.

Consider an SIR disease in a well-mixed population having  $N$  individuals and a given  $\mathcal{R}_0$ . Let  $v$  be a randomly chosen individual from the directed graph created by placing edges from each node to all those nodes they would transmit to if infected.

- Show that if the average number of offspring is  $\mathcal{R}_0$ , then so is the average number of infectors.
- If there are exactly  $\mathcal{R}_0 N$  edges in the directed graph and each recipient is chosen uniformly at random from the population (independent of any previous choice), argue that the number of transmissions  $v$  receives has a binomial distribution with  $\mathcal{R}_0 N$  trials and probability  $\mathcal{R}_0/N$ . (technically we must allow edges from  $v$  to  $v$ )
- Argue that if  $\mathcal{R}_0$  remains fixed as  $N \rightarrow \infty$ , then the number of transmissions  $v$  receives is Poisson distributed with mean  $\mathcal{R}_0$ .

### Solution 4.1

- From Graph Theory, the average in-degree must equal the average out-degree.



Alternately we can note that the average number of edges coming out of a node is the total number of edges divided by  $N$ , so the total number of edges is expected to be  $\mathcal{R}_0 N$ . Then the average number of edges in is also the total number divided by  $N$ , that is,  $\mathcal{R}_0$ .

- b. Individual  $v$  receives a given transmission with probability  $1/N$ . There are  $\mathcal{R}_0 N$  edges, each of which could reach  $v$  with probability  $1/N$ . So this defines a binomial distribution.
- c. As  $N$  increases, the probability a particular edge goes to  $v$  is  $1/N$ , while the number of edges is  $\mathcal{R}_0 N$ . This defines a Poisson distribution with mean  $\mathcal{R}_0$ .

**Exercise 4.2** Explain why for large  $N$  the probability  $v$  is still susceptible at generation  $g$  if she was initially susceptible is  $\chi(S(g-1)/N)$ .

**Solution 4.2** Consider the directed graph created by placing an edge from a node to each node it would infect if given the opportunity.  $v$  is susceptible at generation  $g$  iff  $v$  was initially susceptible, and no ancestor of  $v$  was infected by generation  $g-1$ .

The probability a randomly chosen ancestor was susceptible at generation  $g-1$  is  $S(g-1)/N$ . If  $p_i$  is the probability of having  $i$  ancestors, then the probability of being susceptible at generation  $g$  is

$$\sum_i p_i \left( \frac{S(g-1)}{N} \right)^i$$

but this is just  $\chi(S(g-1)/N)$ .

**Exercise 4.3** Use Theorem 4.2 to derive a result like Theorem 4.1, but with nonzero  $\rho$ .

**Solution 4.3** We have  $\lim_{g \rightarrow \infty} S(g) = \lim_{g \rightarrow \infty} (1-\rho)N\chi(S(g-1)/N)$  or  $S(\infty) = (1-\rho)N\chi(S(\infty)/N)$ . Substituting  $r(\infty) = (N - S(\infty))/N$  yields

$$r(\infty) = 1 - (1-\rho)\chi(1-r(\infty))$$

**Exercise 4.4 Final size relations**

Consider the continuous time SIR dynamics as given in System (35)

- a. Assume  $\kappa = 1$  for all individuals, and write down the corresponding equations for  $S$ ,  $I$ ,  $R$ , and  $\theta$ .
- b. At large time  $I \rightarrow 0$ , so  $S(\infty) = N - R(\infty)$ . But also  $S(\infty) = S(0)\psi(\theta(\infty))$ . By writing  $\theta(\infty)$  in terms of  $R(\infty)$ , derive a recurrence relation for  $r(\infty) = R(\infty)/N$  in terms of  $r(\infty)$  and  $\mathcal{R}_0 = \beta/\gamma$ .
- c. Comment on the relation between your result and Theorem 4.1

**Solution 4.4**

- a.  $\psi(x)$  is simply  $x$  and  $\langle K \rangle = 1$ . We have

$$\begin{aligned} S &= N(1-\rho)\theta \\ I &= N \left( 1 - (1-\rho)\theta + \frac{\gamma}{\beta} \ln \theta \right) \\ R &= -\frac{\gamma N}{\beta} \ln \theta \\ \dot{\theta} &= -\frac{\beta I}{N} \theta \end{aligned}$$

- b. Exponentiating the equation for  $R$  yields

$$\begin{aligned} \theta(\infty) &= \exp(-\beta R(\infty)/\gamma N) \\ &= \exp(-\mathcal{R}_0 r(\infty)) \end{aligned}$$

So we have  $\theta(\infty)$  in terms of  $r(\infty)$ . Now

$$\begin{aligned} r(\infty) &= 1 - S(\infty)/N \\ &= 1 - N(1 - \rho)\theta(\infty)/N \\ &= 1 - (1 - \rho)\theta(\infty) \\ &= 1 - (1 - \rho)e^{-\mathcal{R}_0 r(\infty)} \end{aligned}$$

c. This generalizes the result of Theorem 4.1 for arbitrary  $\rho$ .

#### Exercise 4.5 Other relations

- Using the equations from Exercise 4.4, derive the peak prevalence relation, an expression for the maximum value of  $I$ . [at the maximum  $\dot{I} = 0$ , so we start by finding  $\theta$  so that  $\dot{S} + \dot{R} = 0$ .]
- Similarly, find the peak incidence relation, an expression for the maximum rate at which infections occur,  $-\dot{S}$ .

#### Solution 4.5

a. If  $\dot{S} + \dot{R} = 0$ , then

$$N(1 - \rho)\dot{\theta} - \frac{\gamma N}{\beta} \frac{1}{\theta} \dot{\theta} = 0$$

Factoring out  $N\dot{\theta}$  and taking  $\mathcal{R}_0 = \beta/\gamma$ , we are left

$$(1 - \rho) - \frac{1}{\mathcal{R}_0 \theta} = 0$$

So

$$\theta = \frac{1}{\mathcal{R}_0(1 - \rho)}$$

Then peak prevalence is

$$N \left( 1 - \frac{1}{\mathcal{R}_0} + \frac{1}{\mathcal{R}_0} \ln \frac{1}{\mathcal{R}_0(1 - \rho)} \right) = N \left( 1 - \frac{1}{\mathcal{R}_0} - \frac{1}{\mathcal{R}_0} \ln \mathcal{R}_0(1 - \rho) \right)$$

b. At peak incidence  $-\dot{S}$  takes a maximum. That is,  $-\dot{\theta}$  is a maximum. So we need  $\frac{d}{dt} I\theta = 0$ . That is,

$$\begin{aligned} 0 &= \frac{d}{dt} \left( \theta - (1 - \rho)\theta^2 + \frac{\theta}{\mathcal{R}_0} \ln \theta \right) \\ &= \dot{\theta} - (1 - \rho)2\theta\dot{\theta} + \frac{\dot{\theta}}{\mathcal{R}_0} \ln \theta + \frac{1}{\mathcal{R}_0} \dot{\theta} \end{aligned}$$

Factoring out  $\dot{\theta}$  yields

$$0 = 1 - 2(1 - \rho)\theta + \frac{\ln \theta}{\mathcal{R}_0} + \frac{1}{\mathcal{R}_0}$$

So

$$\theta = \frac{1 + (\ln \theta + 1)/\mathcal{R}_0}{2(1 - \rho)}$$

This needs to be solved numerically. Then plugging this result for  $\theta$  in  $-\dot{S} = -(1 - \rho)\beta N \left( 1 - (1 - \rho)\theta + \frac{1}{\mathcal{R}_0} \ln \theta \right) \theta$  yields the peak incidence.

#### Exercise 4.6 Alternate derivation of $s_u$ .

If the rate of transmissions to  $u$  is  $\beta I \kappa_u / N \langle K \rangle$ , then the expected number of transmissions  $u$  has received

is  $\beta\kappa_u \int_0^t I(\tau) d\tau / N \langle K \rangle$  and this is Poisson distributed.

- Let  $f_u(x)$  be the PGF for the number of transmissions  $u$  has received. Find an expression for  $f_u(x)$  in terms of the integral  $\int_0^t I(\tau) d\tau$ .
- Explain why  $f_u(0)$  is the probability  $u$  is still susceptible.
- Find  $f_u(0)$ .

#### Solution 4.6

- $f$  is the PGF for the Poisson distribution

$$f_u(x) = \exp\left(\beta\kappa_u \frac{\int_0^t I(\tau) d\tau}{N \langle K \rangle} (x - 1)\right)$$

- The probability of still be susceptible is the probability of having received 0 transmissions, which in turn is the coefficient of  $x^0$  in the series expansion of  $f_u(x)$ . This is found by setting  $x = 0$ .
- $f_u(0) = \exp\left(-\beta\kappa_u \frac{\int_0^t I(\tau) d\tau}{N \langle K \rangle}\right)$

#### Exercise 4.7 Alternate derivation of Theorem 4.3 in the homogeneous case.

The usual homogeneous SIR equations are

$$\begin{aligned}\dot{S} &= -\beta IS/N \\ \dot{I} &= \beta IS/N - \gamma I \\ \dot{R} &= \gamma I\end{aligned}$$

We will derive system (35) for fixed  $\kappa = 1$  from this system through the use of an integrating factor. Set  $\theta = e^{-\beta \int_0^t I(\tau) d\tau / N}$ .

- Show that  $\dot{\theta} = -\beta I\theta/N$  and so  $\dot{\theta}/\theta = -\beta\dot{R}/N\gamma$ .
- Using the equation for  $\dot{S}$  add  $\beta IS/N$  to both sides and then divide by (the factor  $1/\theta$  is an integrating factor). Show that the expression on the left hand side is  $\frac{d}{dt} S/\theta = 0$  and so

$$\frac{d}{dt} S/\theta = 0.$$

- Solve for  $R$  in terms of  $\theta$ .
- Solve for  $S$  in terms of  $\theta$ .
- Solve for  $I$  in terms of  $\theta$  using  $S + I + R = N$ .

This equivalence was found in [35] and [20].

#### Solution 4.7

- By direct observation

$$\dot{\theta} = -\frac{\beta I(t)}{N} e^{-\beta \int_0^t I(\tau) d\tau / N} = -\beta I\theta/N$$

and dividing by  $\theta$  and substituting  $\dot{R} = \gamma I$  gives

$$\dot{\theta}/\theta = -\beta\dot{R}/N\gamma$$

b. The operations yield

$$\dot{S}/\theta + \beta IS/N\theta = 0$$

So

$$\begin{aligned} \frac{d}{dt} S/\theta &= \frac{\dot{S}}{\theta} - S\dot{\theta}/\theta^2 \\ &= -\beta IS/N\theta - S\beta I\theta/N\theta^2 \\ &= 0 \end{aligned}$$

c. Since  $\dot{R} = \frac{N\gamma}{\beta}\dot{\theta}/\theta$  we have  $R = R(0) + \frac{N\gamma}{\beta} \ln \theta$

d. Since  $\frac{d}{dt} S/\theta = 0$ , we have  $S = S(0)\theta$ .

e. So  $I = N - S(0)\theta - R(0) - \frac{N\gamma}{\beta} \ln \theta$ .

### Exercise 4.8 Alternate derivation of Theorem 4.3.

Consider now a population having many subgroups of susceptibles denoted by  $\kappa$  with the group  $\kappa$  receiving transmissions at rate  $\beta\kappa I/N$  per individual. Once infected, each individual transmits with rate  $\beta\langle K \rangle$  and recovers with rate  $\gamma$ . These assumptions lead to

$$\begin{aligned} \dot{S}_\kappa &= -\beta\kappa \frac{I}{N\langle K \rangle} S_\kappa \\ \dot{I} &= -\gamma I + \beta \frac{I}{N\langle K \rangle} \sum_\kappa \kappa S_\kappa \\ \dot{R} &= \gamma I \end{aligned}$$

Following Exercise 4.7, set  $\theta = e^{-\beta \int_0^t I(\tau) d\tau/N}$  and derive system (35) from these equations by use of an integrating factor.

**Solution 4.8** We write

$$\dot{S}_\kappa + \beta\kappa \frac{I}{N\langle K \rangle} S_\kappa = 0$$

and multiply by  $1/\theta^\kappa$ . Then

$$\frac{d}{dt} \frac{S_\kappa}{\theta^\kappa} = 0$$

and

$$S_\kappa(t) = S_\kappa(0)\theta^\kappa$$

We assume that  $\psi(x) = \sum_\kappa P(\kappa)x^\kappa$ , where  $P(\kappa) = N_\kappa/N$ . Starting with a fraction  $1 - \rho$  randomly infected yields our expression for  $S = \sum_\kappa S_\kappa(t)$ .

Since  $\dot{R} = \gamma I$  and  $\dot{\theta} = -\beta I\theta/N$ , we can substitute to find  $\dot{R} = -\gamma N\dot{\theta}/\theta\beta$ . This can be integrated to find  $R$  in terms of  $\theta$ .

Then simply set  $I = N - S - R$  to complete the calculation.

## 5 Multitype populations

**Exercise 5.1** Consider a vector-borne disease for which each infected individual infects a Poisson-distributed number of vectors, with mean  $\lambda$ . Each infected vector causes  $i$  infections with probability  $p_i = \pi^i(1 - \pi)$  for some  $\pi \in [0, 1]$ . This scenario corresponds to human infection lasting for a fixed time with some constant transmission rate to vectors, and each vector having probability  $\pi$  of living to bite again after each bite and transmitting with probability 1 if biting.

- a. Let  $\alpha_{g|1}$  and  $\alpha_{g|2}$  be the probability that an outbreak would go extinct in  $g$  generations starting with an infected human or vector respectively. Find the vector-valued function  $\vec{\psi}(\vec{x}) = (\psi_1(\vec{x}), \psi_2(\vec{x}))$ . That is, what are the PGFs  $\psi_1(x_1, x_2)$  and  $\psi_2(x_1, x_2)$ ?
- b. Set  $\lambda = 3$  and  $\pi = 0.5$ . Find the probability of an epidemic if one infected human is introduced or if one infected vector is introduced.
- c. For the same values, find the probability of an epidemic if one infected vector is introduced.
- d. Find  $\psi_2(\psi_1(0, x), 0)$ . How should we interpret the terms of its Taylor Series expansion?

**Solution 5.1**

- a. •  $\psi_1(x_1, x_2) = e^{\lambda(x_2-1)}$  (independent of  $x_1$ )  
 •  $\psi_2(x_1, x_2) = \frac{1-\pi}{1-\pi x_1}$  (independent of  $x_2$ )
- b.  $\vec{\psi}(\vec{x}) = \left( e^{3(x_2-1)}, \frac{1}{2-x_1} \right)$ . Starting with  $\vec{x} = (0, 0)$ , and iterating until convergence we have  $\vec{x} = (0.2868, 0.5837)$ . So the epidemic probability is 0.7132 starting from a human and 0.4163 starting from a vector.

**Exercise 5.2** Starting from the equations

$$\begin{aligned}\dot{S}_i &= -\frac{S_i}{N_i} \sum_j \beta_{ij} I_j \\ \dot{I}_i &= -\gamma_i I_i + \frac{S_i}{N_i} \sum_j \beta_{ij} I_j \\ \dot{R}_i &= \gamma_i I_i\end{aligned}$$

use integrating factors to derive System (37).

**Solution 5.2**

- a. We have

$$\begin{aligned}\dot{S}_i &= -\frac{S_i}{N_i} \sum_j \beta_{ij} I_j \\ \dot{S}_i + \left( \frac{\sum_j \beta_{ij} I_j}{N_i} \right) S_i &= 0 \\ \frac{d}{dt} S_i e^{\xi_i} &= 0 \\ S_i &= S_i(0) e^{-\xi_i}\end{aligned}$$

where  $\dot{\xi}_i = \sum_j \beta_{ij} I_j / N_i$ .

It is straightforward to add  $\dot{R}_i = \gamma_i I_i$  and  $I_i = N_i - S_i - R_i$ .

**Exercise 5.3** Assume the population is grouped into subgroups of size  $N_i$  with  $N = \sum_i N_i$  and the  $i$ -th subgroup has a parameter  $\kappa_i$  representing their rate of contact with others. Take

$$\beta_{ji} = \kappa_j \frac{\kappa_i N_i}{\sum_\ell N_\ell \kappa_\ell} \beta$$

to be the transmission rate from type  $i$  individuals to a single type  $j$  individual, and assume all infected individuals recover with the same rate  $\gamma$ .

Define  $\theta = e^{-\beta(\sum_j \kappa_j \int_0^t I_j(\tau) d\tau) / \sum_j \kappa_j N_j}$  and define the PGF  $\psi(x) = \sum_i \frac{N_i}{N} x^i$ . Let  $S = \sum_i S_i$ ,  $I =$

$\sum_i I_i$ , and  $R = \sum R_i$ .

- Explain what assumptions this model makes about interactions between individuals in group  $i$  and  $j$ .
- Show that

$$\begin{aligned} S &= N\psi(\theta) \\ I &= N - S - R \\ \dot{R} &= \gamma I \\ \dot{\theta} &= -\beta\theta \frac{\sum_j \kappa_j I_j}{\sum_j \kappa_j N_j} \end{aligned}$$

with  $\theta(0) = 1$ .

- Explain why  $\frac{\sum_j \kappa_j I_j}{\sum_j \kappa_j N_j} = 1 - \frac{\sum_j \kappa_j S_j}{\sum_j \kappa_j N_j} - \frac{\sum_j \kappa_j R_j}{\sum_j \kappa_j N_j}$ .
- Show that  $\frac{\sum_j \kappa_j S_j}{\sum_j \kappa_j N_j} = \frac{\theta\psi'(\theta)}{\psi'(1)}$ .
- Show that  $\frac{d}{dt} \frac{\sum_j \kappa_j R_j}{\sum_j \kappa_j N_j} = -(\gamma/\beta) \frac{\dot{\theta}}{\theta}$ , and solve for  $\frac{\sum_j \kappa_j R_j}{\sum_j \kappa_j N_j}$  in terms of  $\theta$  assuming  $R_j = 0$  for all  $j$ .
- Thus conclude that

$$\dot{\theta} = -\beta\theta + \beta \frac{\theta^2 \psi'(\theta)}{\psi'(1)} - \theta\gamma \ln \theta$$

### Solution 5.3

- People in group  $i$  interact with others at rate  $\kappa_i$ , and they choose their partners at random at each time from the entire population with probability proportional to the partner's interaction rate.
- This can be derived directly from an integrating factor or by simply substituting in and checking that the ODEs are satisfied.
- If we move the  $S$  and  $R$  terms to the right hand side, the numerator becomes  $\sum_j \kappa_j (I_j + S_j + R_j) = \sum_j \kappa_j N_j$ .
- We have  $S_j = N_j \theta^{\kappa_j}$  So

$$\begin{aligned} \kappa_j S_j &= \kappa_j \theta^{\kappa_j} \\ &= \theta \frac{d}{d\theta} \theta^{\kappa_j} \end{aligned}$$

Thus summing the terms in the numerator gives  $\theta\psi'(\theta)$ . The denominator can be found similarly.

- If we substitute  $\gamma I_j$  for  $\dot{R}_j$  and compare with the  $\dot{\theta}$  equation the ODE is shown. Then integrating both sides gives  $\sum_j \kappa_j R_j / \sum_j \kappa_j N_j = (\gamma/\beta) \ln \theta$ .
- We simply substitute the results of **d** and **e** into **c**. Then substitute this into the  $\dot{\theta}$  equation from **b**.

## A Important properties of PGFs

**Exercise A.1** Prove Property A.2 [write out the sum and show that the derivatives eliminate any  $r_m$  for  $m < n$ , the leading coefficient of the result is  $n!r_n$ , and the later terms are all zero].

**Solution A.1** We first find

$$\begin{aligned} \left(\frac{d}{dx}\right)^n f(x) &= \left(\frac{d}{dx}\right)^n \sum_m r_m x^m \\ &= \sum_m \left(\frac{d}{dx}\right)^n r_m x^m \\ &= \sum_m m(m-1)\cdots(m-n+1)r_m x^{m-n} \end{aligned}$$

- if  $m < n$ , one of the terms in the product  $m(m-1)\cdots(m-n+1)$  is zero.
- if  $m = n$  then  $m(m-1)\cdots(m-n+1)$  becomes  $n!$ , and  $x^{m-n}$  is simply 1.
- if  $m > n$ , then there is a nonzero factor in front and a factor of  $x^{m-n}$ .

When we evaluate this derivative at  $x = 0$ , the terms which are already zero remain zero, the  $m = n$  term remains unchanged, and the other terms all have 0 raised to a positive power so they become zero.

Thus we are left  $r_n$ .

**Exercise A.2 Verification of Equation (38):**

In this exercise we show that the formula in Equation (38) yields  $r_n$ . Assume that the integral is performed on a circle of radius  $R \leq 1$  about the origin.

- a. Write  $f(z) = \sum_m r_m z^m$  and rewrite  $\int_0^1 \frac{f(Re^{2\pi iu})}{R^n e^{2n\pi iu}} du$  as a sum

$$\int_0^1 \frac{f(Re^{2\pi iu})}{R^n e^{2n\pi iu}} du = \sum_m r_m \int_0^1 R^{m-n} e^{2(m-n)\pi iu} du$$

- b. Show that for  $m = n$  the integral in the summation on the right hand side is 1.  
c. Show that for  $m \neq n$ , the integral in the summation on the right hand side is 0.  
d. Thus conclude that the integral on the left hand side must yield  $r_n$ .

**Solution A.2**

- a. Substituting we have

$$\begin{aligned} \int_0^1 \frac{f(Re^{2\pi iu})}{R^n e^{2n\pi iu}} du &= \int_0^1 \frac{\sum_m r_m (Re^{2\pi iu})^m}{R^n e^{2n\pi iu}} du \\ &= \sum_m r_m \int_0^1 R^{m-n} e^{2(m-n)\pi iu} du \end{aligned}$$

- b. For  $m = n$ , both  $R^{m-n}$  and  $e^{2(m-n)\pi iu}$  become 1. Thus we have  $\int_0^1 1 du = 1$ .

c. For  $m \neq n$ , the integral becomes

$$\begin{aligned} \int_0^1 R^{m-n} e^{2(m-n)\pi i u} du &= R^{m-n} \int_0^1 e^{2(m-n)\pi i u} du \\ &= R^{m-n} \int_0^1 e^{2(m-n)\pi i u} du \\ &= R^{m-n} \frac{1}{2(m-n)\pi i} \left( e^{2(m-n)\pi i} - e^{2(m-n)\pi i 0} \right) \\ &= R^{m-n} \frac{1}{2(m-n)\pi i} (1 - 1) \\ &= 0 \end{aligned}$$

where in the last step we use the fact that  $e^{2N\pi i} = \cos 2N\pi + i \sin 2N\pi = 1$  for any integer  $N$ .

d. So the integral  $\int_0^1 \frac{f(Re^{2\pi i u})}{R^n e^{2n\pi i u}} du$  becomes

$$\begin{aligned} \int_0^1 \frac{f(Re^{2\pi i u})}{R^n e^{2n\pi i u}} du &= 0r_0 + 0r_1 + \cdots + 0r_{n-1} + 1r_n + 0r_{n+1} + \cdots \\ &= r_n \end{aligned}$$

**Exercise A.3** Let  $f(z) = e^z = 1 + z + z^2/2 + z^3/6 + z^4/24 + z^5/120 + \cdots$ . Write a program that estimates  $r_0, r_1, \dots, r_5$  using Equation (39) with  $R = 1$ . Report the values to four significant figures for

- $M = 2$
- $M = 4$
- $M = 5$
- $M = 10$
- $M = 20$ .
- How fast is convergence for different  $r_n$ ?

**Solution A.3** The script `integral_exercise.py` which is given as a supplementary file performs these calculations.

- $M = 2$ :

$$\begin{aligned} r_0 &\approx 1.543 \\ r_1 &\approx 1.175 \\ r_2 &\approx 1.543 \\ r_3 &\approx 1.175 \\ r_4 &\approx 1.543 \\ r_5 &\approx 1.175 \end{aligned}$$



b.  $M = 4$ :

$$\begin{aligned}r_0 &\approx 1.042 \\r_1 &\approx 1.008 \\r_2 &\approx 0.5013 \\r_3 &\approx 0.1669 \\r_4 &\approx 1.042 \\r_5 &\approx 1.008\end{aligned}$$

c.  $M = 5$ :

$$\begin{aligned}r_0 &\approx 1.008 \\r_1 &\approx 1.001 \\r_2 &\approx 0.5002 \\r_3 &\approx 0.1667 \\r_4 &\approx 0.04167 \\r_5 &\approx 1.008\end{aligned}$$

d.  $M = 10$ :

$$\begin{aligned}r_0 &\approx 1.000 \\r_1 &\approx 1.000 \\r_2 &\approx 0.500 \\r_3 &\approx 0.1667 \\r_4 &\approx 0.04167 \\r_5 &\approx 0.008333\end{aligned}$$

e.  $M = 20$ :

$$\begin{aligned}r_0 &\approx 1.000 \\r_1 &\approx 1.000 \\r_2 &\approx 0.5000 \\r_3 &\approx 0.1667 \\r_4 &\approx 0.04167 \\r_5 &\approx 0.008333\end{aligned}$$

f. Convergence is quite fast, and predicts  $r_n$  quite well once  $M > n$ .

**Exercise A.4** The derivation in example A.1 was based on looking at what happened after a single flip and then looking  $g - 1$  flips into the future in the inductive step. Derive  $\alpha_g = f(\alpha_{g-1})$  by instead looking  $g - 1$  flips into the future and then considering one additional step. [the distinction between this argument and the previous one becomes useful in the continuous-time case where we use the ‘backward’ or ‘forward’ Kolmogorov equations.]

**Solution A.4** As in the example, we take  $\alpha_g$  to be the probability of failure within the first  $g$  flips, with  $\alpha_0 = 0$  and  $\alpha_1 = 1 - p = f(0)$ .

If the first flip does not come up as “failure”, then the probability that failure occurs within the following  $g - 1$  flips is (by definition)  $\alpha_{g-1}$ . So the probability of failure within the first  $g$  flips is the probability of

failure in the first flip plus the probability of success time  $\alpha_{g-1}$ . That is:

$$\alpha_g = (1 - p) + p\alpha_{g-1} = f(\alpha_{g-1})$$

**Exercise A.5** Consider a fair six-sided die with numbers  $0, 1, \dots, 5$ , rather than the usual  $1, \dots, 6$ . We roll the die once. Then we look at the result, and roll that many copies (if zero, we stop), then we look at the sum of the result and repeat. Define

$$f(x) = \frac{1 + x + \dots + x^5}{6} = \begin{cases} \frac{x^6 - 1}{6(x-1)} & x \neq 1 \\ 1 & x = 1 \end{cases}$$

Define  $\alpha_g$  to be the probability the process stops after  $g$  iterations (with  $\alpha_0 = 0$  and  $\alpha_1 = 1/6$ ).

- Find an expression for  $\alpha_g$ , the probability that by the  $g$ 'th iteration the process has stopped, in terms of  $f(x)$ .
- Rephrase this question in terms of the extinction probability for an infectious disease.

**Solution A.5**

- The probability of dying out after  $g$  iterations is the sum over all  $i$  of the probability that the first roll is an  $i$  and the process dies out after  $g - 1$  iterations. By thinking of each die that rolls an  $i$  as having  $i$  "offspring" we can assign each individual a collection of descendants. The process goes extinct after  $g$  iterations if all the dice in the second roll have no offspring after  $g - 1$  iterations.

$$\sum_{i=0}^5 \frac{1}{6} \alpha_{g-1}^i = f(\alpha_{g-1}) = f^{[g]}(0)$$

- This is equivalent to a disease for which each individual causes 0, 1, 2, 3, 4, or 5 new infections with equal probability. So the extinction probability for such a disease is  $f^{[g]}(0)$ .

**Exercise A.6** Note that if we interchange  $p$  and  $q$  in the PGF of the negative binomial distribution in Table 1, it is simply the PGF of the geometric distribution raised to the power  $\hat{r}$ . A number chosen from the negative binomial can be defined as the number of successful trials (each with success probability  $p$ ) before the  $\hat{r}$ th failure.

Using this and Property A.8, derive the PGF of the negative binomial.

**Solution A.6** We re-express the negative binomial by interchanging what we count as a failure and success. We seek the number of failures that occur before the  $\hat{r}$ th success. We say a success occurs with probability  $\hat{p}$  and failure with probability  $\hat{q} = 1 - \hat{p}$  (which satisfy  $\hat{p} = q$ ,  $\hat{q} = p$ ).

The result for the geometric distribution says that the number of failures between each success has PGF  $\hat{p}/(1 - \hat{q}x)$ . We need the sum of  $\hat{r}$  of these. So by Property A.8, we have that the negative binomial has PGF  $[\hat{p}/(1 - \hat{q}x)]^{\hat{r}}$ . Replacing  $\hat{p}$  by  $q$  and  $\hat{q}$  by  $p$  completes the result.

**Exercise A.7 Sicherman dice** [18, 17].

To motivate this exercise consider two tetrahedral dice, numbered 1, 2, 3, 4. When we roll them we get sums from 2 to 8, each with its own probability, which we can infer from this table:

	2	3	4	5
	3	4	5	6
	4	5	6	7
	5	6	7	8

However another pair of tetrahedral dice, labelled 1, 2, 2, 3 and 1, 3, 3, 5 yields the same sums with the same

probabilities:

	2	3	3	4
	4	5	5	6
	4	5	5	6
	6	7	7	8

We now try to find a similar pair for 6-sided dice. First consider a pair of standard 6-sided dice.

- a. Show that the PGF of each die is  $f(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)/6$ .
- b. Fill in the tables showing the possible sums from rolling two dice (fill in each square with the sum of the two entries) and multiplication for two polynomials (fill in each square with the product of the two entries):


	$x^1$	$x^2$	$x^3$	$x^4$	$x^5$	$x^6$
$x^1$						
$x^2$						
$x^3$						
$x^4$						
$x^5$						
$x^6$						

- c. Explain the similarity.
- d. Show that each step of the following factorization is correct:

$$\begin{aligned}
 f(x) &= \frac{x(1 + x + x^2 + x^3 + x^4 + x^5)}{6} \\
 &= \frac{x(1 + x + x^2)(1 + x^3)}{6} \\
 &= \frac{x(1 + x + x^2)(1 + x)(1 - x + x^2)}{6}.
 \end{aligned}$$

This cannot be factored further, and indeed it can be shown that a property similar to prime numbers holds. Namely, any factorization of  $f(x)f(x)$  as  $h_1(x)h_2(x)$  has the property that each of  $h_1$  and  $h_2$  can be factored into some powers of these “prime” polynomials times a constant.

We seek two new six-sided dice (each different) such that the sum of a roll of the two dice has the same probabilities as the normal dice. The two dice have positive integer values on them (so no fair adding a constant  $c$  to everything on one die and subtracting  $c$  on the other). Let  $h_1(x)$  and  $h_2(x)$  be their PGFs.

- e. Explain why we must have  $h_1(x)h_2(x) = [f(x)]^2$ .
- f. If the dice have numbers  $a_1, \dots, a_6$  and  $b_1, \dots, b_6$ , show that their PGFs are of the form  $h_1(x) = \sum_i x^{a_i}/6$  and  $h_2(x) = \sum_i x^{b_i}/6$  where all  $a_i$  and  $b_i$  are positive integers.
- g. Given the properties we want for the dice, find  $h_1(0)$  and  $h_2(0)$ .

- h. Given the properties we want for the dice, find  $h_1(1)$  and  $h_2(1)$ .
- i. Using the values at  $x = 0$  and  $x = 1$ , explain why  $h_1(x) = x(1 + x + x^2)(1 + x)(1 - x + x^2)^b/6$  and  $h_2(x) = x(1 + x + x^2)(1 + x)(1 - x + x^2)^{2-b}/6$  where  $b$  is 0, 1, or 2.
- j. The case  $b = 1$  gives the normal dice. Consider  $b = 0$  ( $b = 2$  gives the same final result). Find  $h_1(x)$ .  
 $h_2(x) = \frac{1}{6}(x + x^3 + x^4 + x^5 + x^6 + x^8)$
- k. Create the table for the two dice corresponding to  $h_1(x)$  and  $h_2(x)$  and verify that the sums occur with


the same frequency as a normal pair:

**Solution A.7**

- a. Because each  $i$  in  $1, 2, \dots, 6$  has probability  $1/6$ , we have  $f(x) = \sum_{i=1}^6 (1/6)x^i$ .

b.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

	$x^1$	$x^2$	$x^3$	$x^4$	$x^5$	$x^6$
$x^1$	$x^2$	$x^3$	$x^4$	$x^5$	$x^6$	$x^7$
$x^2$	$x^3$	$x^4$	$x^5$	$x^6$	$x^7$	$x^8$
$x^3$	$x^4$	$x^5$	$x^6$	$x^7$	$x^8$	$x^9$
$x^4$	$x^5$	$x^6$	$x^7$	$x^8$	$x^9$	$x^{10}$
$x^5$	$x^6$	$x^7$	$x^8$	$x^9$	$x^{10}$	$x^{11}$
$x^6$	$x^7$	$x^8$	$x^9$	$x^{10}$	$x^{11}$	$x^{12}$

- c. Multiplication of two powers of  $x$  results in addition of the exponents.
- d. We have

$$x(1 + x + x^2)(1 + x^3) = x(1 + x + x^2 + x^3 + x^4 + x^5) = x + x^2 + x^3 + x^4 + x^5 + x^6$$

and

$$(1 + x)(1 - x + x^2) = 1 - x + x^2 + x - x^2 + x^3 = 1 + x^3$$

- e. Since the PGF of the sum of the two dice is equal to the product of the PGFs of the two dice, we must have  $h_1(x)h_2(x) = [f(x)]^2$ .
- f. For each die, the probabilities of the sides are  $1/6$ . So they take the form  $\sum_i x^{a_i}/6$  and  $\sum_i x^{b_i}/6$ .
- g. Because each  $a_i$  and  $b_i$  is positive  $\sum_i 0^{a_i}/6$  and  $\sum_i 0^{b_i}/6$  are both 0.
- h.  $\sum_i 1^{a_i}/6 = 1$  and  $\sum_i 1^{b_i}/6 = 1$ .

i. Because  $h_1(0) = h_2(0) = 0$ , both of them must have a factor of  $x$ . Because  $h_1(1) = h_2(1) = 1$  and the denominators are 6, the numerators when  $x = 1$  must be 6. So there must be a factor of 2 and 3 in the numerator. This requires that each has a single factor of  $(1 + x + x^2)$  and  $(1 + x)$ . Because  $1 - x + x^2$  is 1 when  $x = 1$ , they could have a factor of  $(1 - x + x^2)^{2-b}$  for  $b = 0, 1, \text{ or } 2$ .

j. Taking  $b = 0$  we have

$$\begin{aligned} h_1(x) &= x(1 + x + x^2)(1 + x)/6 \\ &= \frac{x(1 + 2x + 2x^2 + x^3)}{6} \\ &= \frac{x + x^2 + x^2 + x^3 + x^3 + x^4}{6} \end{aligned}$$

The calculation for  $h_2(x)$  is not asked for in the question, but it is

$$\begin{aligned} h_2(x) &= h_1(x)(1 - x + x^2)(1 - x + x^2) \\ &= h_1(x)(1 - x + x^2 - x + x^2 - x^3 + x^2 - x^3 + x^4) \\ &= \frac{x + 2x^2 + 2x^3 + x^4}{6}(1 - 2x + 3x^2 - 2x^3 + x^4) \\ &= \frac{1}{6}(x - 2x^2 + 3x^3 - 2x^4 + x^5 \\ &\quad + 2x^2 - 4x^3 + 6x^4 - 4x^5 + 2x^6 \\ &\quad + 2x^3 - 4x^4 + 6x^5 - 4x^6 + 2x^7 \\ &\quad + x^4 - 2x^5 + 3x^6 - 2x^7 + x^8) \\ &= \frac{1}{6}(x + x^3 + x^4 + x^5 + x^6 + x^8) \end{aligned}$$

	1	2	2	3	3	4
1	2	3	3	4	4	5
3	4	5	5	6	6	7
k. 4	5	6	6	7	7	8
5	6	7	7	8	8	9
6	7	8	8	9	9	10
8	9	10	10	11	11	12

The sums all occur with the same frequency as a normal pair.

If we didn't calculate  $h_2(x)$  in the previous part, we could have inferred the values from knowing what the first die was and what values would be needed to match the normal pair.

### Exercise A.8 Early-time outbreak dynamics

- a. Consider normal dice. The PGF is  $f(x) = (x + x^2 + x^3 + x^4 + x^5 + x^6)/6$ . Consider the process where we roll a die, take the result  $i$ , and then roll  $i$  other dice and look at their sum. What is the PGF of the resulting sum in terms of  $f$ ?
- b. If an infected individual causes anywhere from 1 to 6 infections, all with equal probability, find the PGF for the number of infections in generation 2 if there is one infection in generation 0. [you can express the result in terms of  $f$ ]
- c. And in generation  $g$  (assuming depletion of susceptibles is unimportant)?

### Solution A.8

- By Property A.8, the PGF is  $f(f(x))$ .
- This is equivalent to the dice-rolling example. It is  $f(f(x))$ .
- More generally for generation  $g$ , we have  $f^{[g]}(x)$ .

### Exercise A.9 Understanding cobweb diagrams

From figure 10 the origin of the term “cobweb” may be unclear. Because of properties of PGFs, the more interesting behavior does not occur for our applications. Here we investigate cobweb diagrams in more detail for non-PGF functions. Since we use  $f(x)$  to denote a PGF, in this exercise we use  $z(x)$  for an arbitrary function.

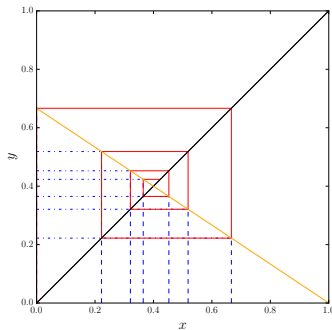
- Consider the line  $z(x) = 2(1 - x)/3$ . Starting with  $x_0 = 0$ , show how the first few iterations of  $x_i = z(x_{i-1})$  can be found using a cobweb diagram (do not explicitly calculate the values).
- Now consider the line  $z(x) = 2(1 - x)$ . The solution to  $z(x) = x$  is  $x = 2/3$ . Starting from an initial  $x_0$  close to (but not quite equal to)  $2/3$ , do several iterations of the cobweb diagram graphically.
- Repeat this with the lines  $z(x) = 1/4 + x/2$  starting at  $x_0 = 0$  and  $z(x) = -1 + 3x$  starting close to where  $x = z(x)$ .
- What is different when the slope is positive or negative?
- Can you predict what condition on the slope’s magnitude leads to convergence to or divergence from the solution to  $x = z(x)$  when  $z$  is a line?

So far we have considered lines  $z(x)$ . Now assume  $z(x)$  is nonlinear and consider the behavior of cobweb diagrams close to a point where  $x = z(x)$ .

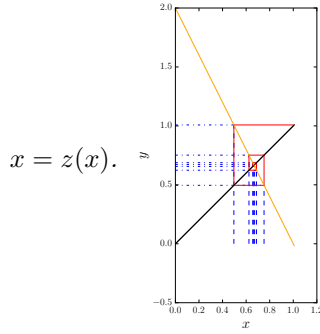
- Use Taylor Series to argue that (except for degenerate cases where  $z'$  is 1 at the intercept) it is only the slope at the intercept that determines the behavior sufficiently close to the intercept.

### Solution A.9

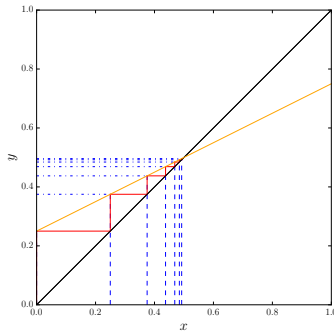
- For  $z(x) = 2(1 - x)/3$  with  $x_0 = 0$ , the cobweb diagram spirals in towards 0.4, the solution to  $x = z(x)$ .



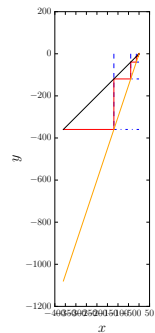
- For  $z(x) = 2(1 - x)$  with  $x_0 = 0.664$ , the cobweb diagram spirals away from  $2/3$ , the solution to



c. For  $z(x) = 1/4 + x/2$ , the cobweb diagram steps in to  $x = 1/2$ , the solution  $x = z(x)$ .



For  $z(x) = -1 + 3x$ , the cobweb diagram steps away from  $x = 1/2$ , the solution  $x = z(x)$ .



- d. When the slope is positive, each successive value is on the same side of the intercept. When it is negative the values alternate.
- e. If we take  $x^*$  to be the solution  $x^* = z(x^*)$ , then when  $|z'(x^*)| > 1$  the values diverge from  $x^*$ , while when  $|z'(x^*)| < 1$  they converge.
- f. Sufficiently close to the intersection where  $x^* = z(x^*)$ , we can locally treat  $z(x)$  as a line. If  $z'(x^*)$  is 1 at this intersection, this line lies on top of  $z(x) = x$  and the correction to the line makes a difference. For all other values the nonlinear terms can be neglected.

**Exercise A.10 Structure of fixed points of  $f(x)$ .**

Consider a PGF  $f(x) = \sum_i r_i x^i$ , and assume  $r_0 > 0$ .

- a. Show that  $f(1) = 1$  and  $f(0) > 0$ .
- b. Show that  $f(x)$  is convex (that is  $f''(x) \geq 0$ ) for  $x > 0$ . [hint  $r_i \geq 0$  for all  $i$ ]
- c. Thus argue that if  $f'(1) \leq 1$ , then  $x = f(x)$  has only one solution to  $x = f(x)$  in  $[0, 1]$ , namely  $f(1) = 1$ . It may help to draw pictures of  $f(x)$  and the function  $y = x$  for  $x$  in  $[0, 1]$ .
- d. Explain why if there is a point  $x_0 \neq 1$  where  $f(x_0) = x_0$  and  $f(x) > x$  for  $x$  in some region  $(x_0, x_1)$

then  $0 < f'(x_0) < 1$ .

- e. Thus show that if  $f'(1) > 1$  then there are exactly two solutions to  $x = f(x)$  in  $[0, 1]$ , one of which is  $x = 1$ .

### Solution A.10

- a.  $f(1) = \sum_i r_i 1^i = \sum_i r_i = 1$  because the  $r_i$  form a probability distribution and  $f(0) = \sum_i r_i 0^i = r_0 > 0$ .
- b.  $f''(x) = \sum_i i(i-1)r_i x^{i-2}$ . Because  $r_i \geq 0$  and  $x \geq 0$  this sum is  $\geq 0$ . [the inequality is strict if any  $r_i > 0$  for  $i \geq 2$ ].
- c. Note first that  $f(1) = 1$ . If  $f'(1) \leq 1$ , then for all  $x$  in  $0 \leq x < 1$ , we have  $f'(1) < 1$ . So because the line  $y = x$  has a larger slope, it lies strictly below  $f(x)$  for all  $x$  in  $[0, 1)$ . [a more rigorous argument can be made using the mean value theorem: assume some  $x_0$  in  $[0, 1)$  has  $f(x_0) = x_0$ . Then by the mean value theorem there is an  $\hat{x} \in (x_0, 1)$  where  $f'(\hat{x})$  is equal to the slope of the line connecting the points  $(x_0, f(x_0))$  and  $(1, f(1))$ . This line is  $y = x$ , and has slope 1, so  $f'(\hat{x}) = 1$ . No such  $\hat{x}$  exists, so the assumption that such an  $x_0$  exists must be false.]
- d.  $f'(x_0) > 0$  because  $f'(x_0) = \sum_i i r_i x_0^{i-1}$  and unless  $r_0 = 1$ , this is strictly positive. If  $r_0 = 1$ , then  $f(x) = 1$  and no such  $x_0$  exists. To show  $f'(x_0) < 1$ , note that for  $f(x)$  to be less than  $x$  for  $x > x_0$ , but equal to  $x$  for  $x = x_0$ , it must have smaller slope than that of  $y = x$  [a more rigorous proof uses the mean value theorem to show that there is an  $\hat{x} \in (x_0, x_1)$  with slope  $< 1$  and so since  $f''(x) > 0$ , we have  $f'(x_0) < f'(\hat{x}) < 1$ ].
- e. Again we start with the observation that  $f(1) = 1$ . Because  $f'(x) > 1$  at  $x = 1$ , there must be a region  $(\hat{x}, 1)$  such that if  $x \in (\hat{x}, 1)$  then  $f(x) < x$  as it has greater slope than the line  $y = x$  [again, a more rigorous proof would use the mean-value theorem, and the fact that  $f'(x)$  is continuous at 1]. However, for the curve  $f(x)$  to reach  $f(0) > 0$  at  $x = 0$ , it must somewhere cross the line [rigorously, we can use the intermediate value theorem applied to the function  $f(x) - x$  with the point  $x = 0$  and some  $x = x_1 \in (x_0, 1)$ ]. At the point  $x_0$  where  $f(x)$  crosses  $y = x$ , our previous result shows that  $f'(x_0) < 1$ . So for  $x \in (0, x_0)$  we have  $f'(x) < 1$  as well. Thus  $f(x)$  cannot cross  $y = x$  at any other point in  $(0, x_0)$  [again, the mean value theorem does this more rigorously: if it did cross again, then we could prove some  $\hat{x} \in (0, x_0)$  where  $f'(\hat{x}) = 1$ ].

### Exercise A.11 Alternate derivation of Equation (40)

An alternate way to derive Equation (40) is through directly calculating  $\dot{r}_i$ .

- a. Explain why  $\dot{r}_i = -\sum_m \lambda_m i r_i + \sum_m \lambda_m (i - m + 1) r_{i-m+1}$ .
- b. Taking  $\dot{f}(x, t) = \sum_i \dot{r}_i x^i$ , derive Equation (40).

### Solution A.11

- a. We simply look at the rate  $r_i$  is decreasing due to events that move the system out of state  $i$  plus the rate it increases due to events that move the system into state  $i$ . The total rate moving out is  $\sum_m \lambda_m i r_i$ . The total rate in is  $\sum_m \lambda_m (i - m + 1) r_{i-m+1}$ .



b.

$$\begin{aligned}
\dot{f}(x, t) &= \sum_i \dot{r}_i x^i \\
&= \sum_i \left( -\sum_m \lambda_m i r_i + \sum_m \lambda_m (i - m + 1) r_{i-m+1} \right) x^i \\
&= -\sum_i \sum_m \lambda_m i r_i x^i + \sum_i \sum_m \lambda_m (i - m + 1) r_{i-m+1} x^i \\
&= -\Lambda \sum_i i r_i x^i + \sum_m \lambda_m \sum_i (i - m + 1) r_{i-m+1} x^i \\
&= -\Lambda x \sum_i i r_i x^{i-1} + \sum_m \lambda_m x^m \sum_i (i - m + 1) r_{i-m+1} x^{i-m} \\
&= -\Lambda x \frac{\partial}{\partial x} \sum_i r_i x^i + \sum_m \lambda_m x^m \frac{\partial}{\partial x} \sum_i r_{i-m+1} x^{i-m+1} \\
&= -\Lambda x \frac{\partial}{\partial x} f(x, t) + \sum_m \lambda_m x^m \frac{\partial}{\partial x} f(x, t) \\
&= -\Lambda x \frac{\partial}{\partial x} f(x, t) + \Lambda h(x) \frac{\partial}{\partial x} f(x, t) \\
&= \Lambda (h(x) - x) \frac{\partial}{\partial x} f(x, t)
\end{aligned}$$

**Exercise A.12** In many cases interactions between two individuals of the same type are important. These may occur with rate  $i(i-1)$  or  $i^2$  depending on the specific details. Assume we have only a single type of individual with PGF  $f(x, t) = \sum_i r_i(t) x^i$ .

- If a collection of events to replace two individuals with  $m$  individuals occur with rate  $\beta_m i(i-1)$ , find how write a PDE for  $f$ . Your final result should contain  $\frac{\partial^2}{\partial x^2} f(x, t)$ . Use  $\mathfrak{B} = \sum_m \beta_m$  and  $g(x) = \sum_m \beta_m x^m / \mathfrak{B}$ . Follow the derivation of Equation (40).
- If instead the events replace two individuals with  $m$  individuals and occur with rate  $\beta_m i^2$ , find how to incorporate them into a PDE for  $f$ . Your final result should contain  $\frac{\partial}{\partial x} (x \frac{\partial}{\partial x} f(x, t))$  or equivalently  $\frac{\partial}{\partial x} f(x, t) + x \frac{\partial^2}{\partial x^2} f(x, t)$ .

**Solution A.12**

- Let  $\beta_m i(i-1)$  denote the rate at which the system goes from a state with  $i$  individuals to  $i-2+m$  individuals (that is, two individuals are replaced by  $m$ ). Then

$$\begin{aligned}
f(x, t + \Delta t) &= \sum_i r_i(t) \left[ \sum_m (\beta_m i(i-1) \Delta t) x^{i+m-2} + \left( 1 - \sum_m \beta_m i(i-1) \Delta t \right) x^i \right] + \mathcal{O}(\Delta t) \\
&= \sum_i r_i x^i + \sum_m \beta_m (\Delta t) (x^m - x^2) \sum_i r_i i(i-1) x^{i-2} + \mathcal{O}(\Delta t) \\
&= f(x, t) + \mathfrak{B} \Delta t [g(x) - x^2] \sum_i r_i i(i-1) x^{i-2} + \mathcal{O}(\Delta t) \\
&= f(x, t) + \mathfrak{B} \Delta t [g(x) - x^2] \frac{\partial^2}{\partial x^2} f(x, t) + \mathcal{O}(\Delta t)
\end{aligned}$$

Plugging this into

$$\frac{\partial}{\partial t} f(x, t) = \lim_{\Delta t \rightarrow 0} \frac{f(x, t + \Delta t) - f(x, t)}{\Delta t}$$

yields

$$\frac{\partial}{\partial t} f(x, t) = \mathfrak{B}[g(x) - x^2] \frac{\partial^2}{\partial x^2} f(x, t)$$

- b. The proof is almost the same as the previous case: Let  $\beta_m i^2$  denote the rate at which the system goes from a state with  $i$  individuals to  $i - 2 + m$  individuals (that is, two individuals are replaced by  $m$ ). Then

$$\begin{aligned} f(x, t + \Delta t) &= \sum_i r_i(t) \left[ \sum_m (\beta_m i^2 \Delta t) x^{i+m-2} + \left( 1 - \sum_m \beta_m i^2 \Delta t \right) x^i \right] + \mathcal{O}(\Delta t) \\ &= \sum_i r_i x^i + \sum_m \beta_m (\Delta t) (x^m - x^2) \sum_i r_i i^2 x^{i-2} + \mathcal{O}(\Delta t) \\ &= f(x, t) + \mathfrak{B} \Delta t [g(x) - x^2] \sum_i r_i i^2 x^{i-2} + \mathcal{O}(\Delta t) \\ &= f(x, t) + \mathfrak{B} \Delta t [g(x) - x^2] \frac{1}{x} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} f(x, t) + \mathcal{O}(\Delta t) \end{aligned}$$

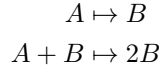
Plugging this into

$$\frac{\partial}{\partial t} f(x, t) = \lim_{\Delta t \rightarrow 0} \frac{f(x, t + \Delta t) - f(x, t)}{\Delta t}$$

yields

$$\frac{\partial}{\partial t} f(x, t) = \mathfrak{B}[g(x) - x^2] \frac{1}{x} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} f(x, t)$$

**Exercise A.13** Consider a chemical system that begins with some initial amount of chemical  $A$ . Let  $i$  denote the number of molecules of species  $A$ . A molecule of  $A$  spontaneously degrades into a molecule of  $B$ , with rate  $\xi$  per molecule. Let  $j$  denote the number of molecules of species  $B$ . Species  $B$  reacts with  $A$  at rate  $\eta j$  to produce new molecules of species  $B$ . The reactions are denoted



Let  $r_{i,j}(t)$  denote the probability of  $i$  molecules of  $A$  and  $j$  molecules of  $B$  at time  $t$ . Let  $f(x, y, t) = r_{i,j}(t) x^i y^j$  be the PGF. Find the Forward Kolmogorov Equation for  $f(x, y, t)$ .

**Solution A.13**

$$\begin{aligned} \frac{\partial}{\partial t} f(x, y, t) &= \xi(x^{-1}y - 1)x \frac{\partial}{\partial x} f(x, y, t) + \eta(x^{-1}y - 1)xy \frac{\partial}{\partial x} \frac{\partial}{\partial y} f(x, y, t) \\ &= \left( \xi(y - x) \frac{\partial}{\partial x} + \eta(y^2 - xy) \frac{\partial}{\partial x} \frac{\partial}{\partial y} \right) f(x, y, t) \end{aligned}$$

**Exercise A.14** In this exercise we generalize Property A.12 for the case where there are two types of individuals  $A$  and  $B$  with counts  $i$  and  $j$ .

Assume events occur spontaneously with rate  $\lambda_{m,n}i$  to remove an individual of type  $A$  and replace it with  $m$  of type  $A$  and  $n$  of type  $B$ , or they occur spontaneously with rate  $\zeta_{m,n}j$  to remove an individual of type  $B$  and replace it with  $m$  of type  $A$  and  $n$  of type  $B$ .

Set  $\Lambda = \sum_{m,n} \lambda_{m,n}$  and  $\mathfrak{Z} = \sum_{m,n} \zeta_{m,n}$ . Let  $f_{1,0}(x, y, t)$  denote the outcome beginning with one individual of type  $A$  and  $f_{0,1}(x, y, t)$  denote the outcome beginning with one individual of type  $B$ .

- Write  $f_{1,0}(x, y, \Delta t)$  and  $f_{0,1}(x, y, \Delta t)$  in terms of  $h(x, y) = \sum_{m,n} \lambda_{m,n} x^m y^n / \Lambda$  and  $g(x, y) = \sum_{m,n} \zeta_{m,n} x^m y^n / \mathfrak{Z}$ .
- Use Property A.8, write  $f_{1,0}(x, \Delta t + t)$  and  $f_{0,1}(x, \Delta t + t)$  in terms of  $f_{1,0}$  and  $f_{0,1}$  evaluated at  $t$  and  $\Delta t$ . The answer should resemble Equation (42).
- Derive expressions for  $\frac{\partial}{\partial t} f_{1,0}(x, y, t)$  and  $\frac{\partial}{\partial t} f_{0,1}(x, y, t)$ .

d. Use this to derive Equation (22).

**Solution A.14**

a.

$$\begin{aligned}
f_{1,0}(x, y, \Delta t) &= \sum_i r_{i,j}(0) x^i \left( 1 - \sum_{m,n} i \lambda_{m,n} \Delta t + \sum_{m,n} i \lambda_{m,n} x^{m-1} y^n \right) + \mathcal{O}(\Delta t) \\
&= x \left( 1 - \sum_{m,n} \lambda_{m,n} \Delta t + \sum_{m,n} \lambda_{m,n} x^{m-1} y^n \right) + \mathcal{O}(\Delta t) \\
&= x - x(\Delta t) \sum_{m,n} \lambda_{m,n} + \sum_{m,n} \lambda_{m,n} x^m y^n + \mathcal{O}(\Delta t) \\
&= x + (\Delta t) \Lambda [h(x, y) - x] + \mathcal{O}(\Delta t)
\end{aligned}$$

Similarly

$$f_{0,1}(x, y, \Delta t) = y + (\Delta t) \mathfrak{Z} [g(x, y) - y] + \mathcal{O}(\Delta t)$$

b. We find

$$f_{1,0}(x, y, \Delta t + t) = f_{1,0}(f_{1,0}(x, y, t), f_{0,1}(x, y, t), \Delta t)$$

and

$$f_{0,1}(x, y, \Delta t + t) = f_{0,1}(f_{1,0}(x, y, t), f_{0,1}(x, y, t), \Delta t)$$

c. We now have

$$\begin{aligned}
\frac{\partial}{\partial t} f_{1,0}(x, y, t) &= \lim_{\Delta t \rightarrow 0} \frac{f_{1,0}(x, y, \Delta t + t) - f_{1,0}(x, y, t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{f_{1,0}(f_{1,0}(x, y, t), f_{0,1}(x, y, t), \Delta t) - f_{1,0}(x, y, t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \frac{f_{1,0}(x, y, t) + (\Delta t) \Lambda [h(f_{1,0}(x, y, t), f_{0,1}(x, y, t)) - f_{1,0}(x, y, t)] + \mathcal{O}(\Delta t) - f_{1,0}(x, y, t)}{\Delta t} \\
&= \Lambda [h(f_{1,0}(x, y, t), f_{0,1}(x, y, t)) - f_{1,0}(x, y, t)] + \lim_{\Delta t \rightarrow 0} \frac{\mathcal{O}(\Delta t)}{\Delta t} \\
&= \Lambda [h(f_{1,0}(x, y, t), f_{0,1}(x, y, t)) - f_{1,0}(x, y, t)]
\end{aligned}$$

Similarly

$$\frac{\partial}{\partial t} f_{0,1}(x, y, t) = \mathfrak{Z} [g(f_{1,0}(x, y, t), f_{0,1}(x, y, t))]$$

d. To derive Equation (22), we take infected individuals to be type A and recovered individuals to be type B. We replace  $x$  with  $\tilde{y}$ , and  $y$  with  $\tilde{z}$ . We will drop the tildes later.

The events that can happen are that a single infected individual can be replaced by 2 infected individuals (with rate  $\beta$  per infected individual) or by 1 recovered individual (with rate  $\gamma$  per infected individual). So  $\Lambda = (\beta + \gamma)$ , and  $h(\tilde{y}, \tilde{z}) = (\beta \tilde{y}^2 + \gamma \tilde{z}) / (\beta + \gamma)$ . There are no events that can happen to recovered individuals, so  $\mathfrak{Z} = 0$  and  $g$  could be anything.

So

$$\begin{aligned}
\frac{\partial}{\partial t} f_{1,0}(\tilde{y}, \tilde{z}, t) &= (\beta + \gamma) \left[ \frac{\beta f_{1,0}(\tilde{y}, \tilde{z}, t)^2 + \gamma f_{0,1}(\tilde{y}, \tilde{z}, t)}{\beta + \gamma} - f_{1,0}(\tilde{y}, \tilde{z}, t) \right] \\
\frac{\partial}{\partial t} f_{0,1}(\tilde{y}, \tilde{z}, t) &= 0
\end{aligned}$$

Since  $f_{0,1}(\tilde{y}, \tilde{z}, t)$  is constant and initially it is simply  $\tilde{z}$ , we conclude that it is always  $\tilde{z}$ . Thus we get

$$\frac{\partial}{\partial t} f_{1,0}(\tilde{y}, \tilde{z}, t) = (\beta + \gamma) \left[ \frac{\beta f_{1,0}(\tilde{y}, \tilde{z}, t)^2 + \gamma \tilde{z}}{\beta + \gamma} - f_{1,0}(\tilde{y}, \tilde{z}, t) \right]$$

Then replacing  $f_{1,0}$  by  $\Pi$ ,  $\tilde{y}$  by  $y$  and  $\tilde{z}$  by  $z$  and replacing  $\frac{\beta f_{1,0}(\tilde{y}, \tilde{z}, t)^2 + \gamma \tilde{z}}{\beta + \gamma}$  by  $\hat{\mu}(\Pi(y, z, t), z)$  completes the result.

## B Proof of Theorems 2.7 and 3.6

**Exercise B.1** If we do not think of an infected individual as disappearing and being replaced by two infected individuals when a transmission happens, but rather, we count up all of the transmissions the individual causes, we get a geometric distribution with  $q = \beta/(\beta + \gamma)$ . The details are in Exercise 3.2. Use this along with Theorem 2.7 and Table 6 (which was derived in exercise 2.13) to give a different proof of Theorem 3.6.

**Solution B.1** The offspring distribution is geometric with  $q = \beta/(\beta + \gamma)$ .

The result in Table 6 predicts  $j$  infections with probability

$$\frac{1}{j} \binom{2j-2}{j-1} \left( \frac{\gamma}{\beta + \gamma} \right)^j \left( \frac{\beta}{\beta + \gamma} \right)^{j-1} = \frac{1}{j} \frac{\gamma^j \beta^{j-1}}{(\beta + \gamma)^{2j-1}} \binom{2j-2}{j-1}$$