

Chapter 3

Applications of Multi-Type Branching Processes

3.1 Introduction

Two applications of multi-type branching processes to epidemic models are presented. The first application is to an SEIR epidemic model and the second application is to the same epidemic model but with dispersal. The SEIR epidemic is modeled as a two-type branching process. Occurrence of an outbreak depends on the number of exposed and infectious individuals. It is shown that the offspring pgfs for the exposed and infectious populations lead to an explicit formula for the probability of an outbreak. In the SEIR model with dispersal, the case of two regions with different healthcare situations are considered. One region has poor healthcare versus another region with excellent healthcare. It is shown that the rate and the direction of movement have a large impact on the occurrence of an outbreak. Branching process theory is used to investigate the probability of an outbreak when the movement rates differ between the two regions.

Although the SIR and SEIR epidemic models are simple, they are often used as a first approximation during or after disease outbreaks to provide estimates of the potential spread of the disease or to understand the pattern of spread. For example, SIR and SEIR epidemic models in conjunction with data provided useful information about the spread of the 2002–2003 SARS (Severe Acute Respiratory Syndrome) pandemic which began in China, the 2009–2010 H1N1 influenza pandemic which began in Mexico, and the 2014 Ebola outbreak in Africa [11, 22, 33].

3.2 SEIR Epidemic

Consider an SEIR epidemic model, where S , E , I , and R are the susceptible, exposed, infectious, and recovered individuals, respectively. With disease-related mortality rate α , the population size is not constant, $S(t) + E(t) + I(t) + R(t) = N(t)$. The deterministic SEIR ODE model has the form:

$$\begin{aligned}
 \frac{dS(t)}{dt} &= -\beta \frac{S(t)}{N(t)} I(t) \\
 \frac{dE(t)}{dt} &= \beta \frac{S(t)}{N(t)} I(t) - \delta E(t) \\
 \frac{dI(t)}{dt} &= \delta E(t) - \gamma I(t) - \alpha I(t) \\
 \frac{dR(t)}{dt} &= \gamma I(t).
 \end{aligned} \tag{3.1}$$

Births, deaths, and temporary immunity are not included in this model. However, the basic reproduction number near the disease-free state has the same form as in the SIRS model considered in Chapter 2 [36]:

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \alpha}. \tag{3.2}$$

For the CTMC SEIR epidemic model, let $\mathbf{X}(t) = (X_1(t), X_2(t), X_3(t), X_4(t))$ denote the discrete random variables for the four states, $(S(t), E(t), I(t), R(t))$. The transition rates for the CTMC SEIR epidemic model (MC Rates) and those for the corresponding branching process approximation (BP rates) for exposed and infectious populations, X_2 and X_3 are given in Table 3.1. Because the rates are nonlinear, the solution of the deterministic model does not represent the mean of the stochastic model. Note that event 1 has a nonlinear transition rate in the MC model but a linear rate in the branching process approximation.

Table 3.1 Transition rates for the CTMC SEIR epidemic model (MC Rates) and for the corresponding branching process approximation for exposed and infectious individuals (BP Rates).

Event	$\Delta \mathbf{X}(t)$	MC Rates	BP Rates
1	$(-1, 1, 0, 0)$	$\beta \frac{X_1(t)}{N(t)} I(t)$	$\beta I(t)$
2	$(0, -1, 1, 0)$	$\delta X_2(t)$	$\delta X_2(t)$
3	$(0, 0, -1, 1)$	$\gamma X_3(t)$	$\gamma X_3(t)$
4	$(0, 0, -1, 0)$	$\alpha X_3(t)$	$\alpha X_3(t)$

To compute the probability of epidemic extinction for the multi-type branching process, the pgfs for the random variables, X_2 and X_3 , are defined. Applying the transition rates from Table 3.1, the pgfs for X_2 and X_3 are

$$f_1(s_1, s_2) = \frac{\delta s_2}{\delta}$$

$$f_2(s_1, s_2) = \frac{\beta s_1 s_2 + \gamma + \alpha}{\beta + \gamma + \alpha}.$$

Although f_1 is a simple function, f_2 is not. The expectation matrix of the pgfs is

$$M = \begin{bmatrix} 0 & 1 \\ \beta & \beta \\ \beta + \gamma + \alpha & \beta + \gamma + \alpha \end{bmatrix}.$$

Matrix $J = \Lambda(M - I)$ is

$$J = \begin{bmatrix} -\delta & \delta \\ \beta & -\gamma - \alpha \end{bmatrix},$$

where $\Lambda = \text{diag}(\delta, \beta + \gamma + \alpha)$. Both matrices are irreducible. It is clear that if $\mathcal{R}_0 > 1$, the branching process is supercritical. In particular, if $\mathcal{R}_0 > 1$, the unique fixed point of the pgfs is $(q_1^*, q_2^*) \in (0, 1)^2$, where $q_i^* = 1/\mathcal{R}_0$ (Whittle's result). The difference between the SIR and SEIR CTMC models is that the exposed period increases the time until extinction and the presence of both exposed and infectious individuals increases the probability of an outbreak. This latter result can be seen in the probability of extinction (no major outbreak) for the CTMC SEIR epidemic model, which is given by

$$\mathbb{P}_0(e_0, i_0) \approx (1/\mathcal{R}_0)^{e_0+i_0}.$$

A numerical example of the SEIR CTMC model along with the deterministic solution is given in Figure 3.1. Three of the four sample paths represent an outbreak, whereas in one there is no outbreak. The probability of a major outbreak is $1 - \mathbb{P}_0(1, 0) = 0.625$.

If additional mortality occurs during the exposed period at rate $\varepsilon X_2(t)$, then the pgf for X_2 is

$$f_1(s_1, s_2) = \frac{\delta s_2 + \varepsilon}{\delta + \varepsilon}.$$

The basic reproduction number for the ODE model with mortality during the exposed period differs from (3.2) and is equal to

$$\mathcal{R}_0 = \frac{\beta \delta}{(\delta + \varepsilon)(\gamma + \alpha)}. \quad (3.3)$$

If $\mathcal{R}_0 > 1$, then the fixed point can be explicitly determined,

$$q_1^* = \frac{\delta}{\delta + \varepsilon} \frac{1}{\mathcal{R}_0} + \frac{\varepsilon}{\delta + \varepsilon}$$

$$q_2^* = \frac{1}{\mathcal{R}_0}.$$

A model with natural births and deaths in all stages yields a similar result [4]. The value for the probability of extinction is greater in the exposed period than in the infectious period, $q_1^* > q_2^*$. This is a reasonable result since in the exposed period, individuals may die with probability $\varepsilon/(\delta + \varepsilon)$ before becoming infectious or become infectious but not transmit the disease with probability $\delta/[(\delta + \varepsilon)\mathcal{R}_0]$.

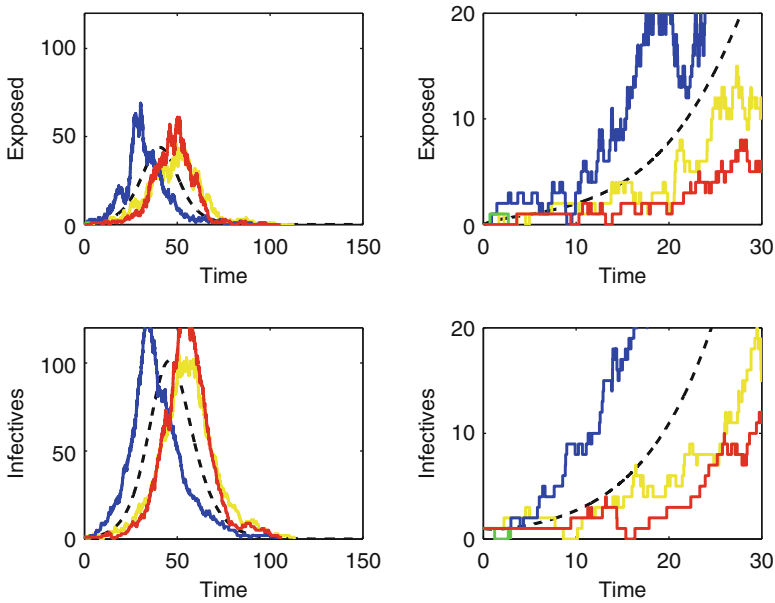


Fig. 3.1 Four sample paths of the CTMC SEIR epidemic model along with the deterministic solution (dashed curve). Parameter values are $\beta = 0.4$, $\delta = 0.4$, $\gamma = 0.1$, $\alpha = 0.05$, and $\varepsilon = 0$. Initial values are $S(0) = 499$, $E(0) = 1$, $I(0) = 0 = R(0)$. The basic reproduction number $\mathcal{R}_0 = 2.67$. The probability of a major outbreak is $1 - \mathbb{P}_0(1, 0) = 0.625$.

3.3 Epidemic Dispersal

Suppose disease is spread between two populations each occupying different regions or patches and modeled by the SEIR epidemic equations within each patch. In population 1, poor healthcare facilities result in frequent disease outbreaks. In population 2, better healthcare facilities and reduced mortality and recovery rates result in no major outbreaks. For population 1, the basic reproduction number is greater than one but for population 2, the basic reproduction number is less than one. With dispersal between these two populations, the outcome changes depending on the direction and the rate of dispersal.

Let the disease parameters for each of these populations be denoted as β_i , δ_i , γ_i , and ε_i , $i = 1, 2$. The rate of dispersal from population 1 to 2 is m_1 and the rate from

population 2 to 1 is m_2 . For simplicity, the model assumes that all individuals within each stage, S , E , I , or R , disperse at the same rates, i.e., with rates m_1 and m_2 . The disease is spread by the movement of exposed or infectious individuals between these two populations. The deterministic model for population 1 with dispersal is

$$\begin{aligned}
 \frac{dS_1(t)}{dt} &= -\beta_1 \frac{S_1(t)}{N_1(t)} I_1(t) - m_1 S_1(t) + m_2 S_2(t) \\
 \frac{dE_1(t)}{dt} &= \beta_1 \frac{S_1(t)}{N_1(t)} I_1(t) - \delta_1 E_1(t) - \varepsilon_1 E_1(t) - m_1 E_1(t) + m_2 E_2(t) \\
 \frac{dI_1(t)}{dt} &= \delta_1 E_1(t) - \gamma_1 I_1(t) - \alpha_1 I_1(t) - m_1 I_1(t) + m_2 I_2(t) \\
 \frac{dR_1(t)}{dt} &= \gamma_1 I_1(t) - m_1 R_1(t) + m_2 R_2(t).
 \end{aligned} \tag{3.4}$$

A similar system holds for population 2. Without dispersal, $m_i = 0$, the basic reproduction number \mathcal{R}_{0i} , $i = 1, 2$, for each population is given by formula (3.3), where the parameters for population 1 or 2 have subscripts 1 or 2, respectively. We assume $\mathcal{R}_{01} > 1$ and $\mathcal{R}_{02} < 1$.

A branching process approximation for the corresponding CTMC SEIR model for two patches can be applied if the population size is large but the exposed and infectious population sizes are small. For the branching process, we are only interested in the exposed and infectious stages. The direction and the rate of movement of individuals in these disease stages have a large impact on the probability of an outbreak.

Let $\mathbf{X}(t) = (X_1(t), X_2(t), X_3(t), X_4(t))$ denote the four discrete random variables for stages E_1 , I_1 , E_2 , and I_2 , respectively. The four probability generating functions of the approximating branching process are

$$\begin{aligned}
 f_1(s_1, s_2, s_3, s_4) &= \frac{\delta_1 s_2 + \varepsilon_1 + m_1 s_3}{\delta_1 + m_1 + \varepsilon_1} \\
 f_2(s_1, s_2, s_3, s_4) &= \frac{\beta_1 s_1 s_2 + \gamma_1 + \alpha_1 + m_1 s_4}{\beta_1 + \gamma_1 + \alpha_1 + m_1} \\
 f_3(s_1, s_2, s_3, s_4) &= \frac{\delta_2 s_4 + \varepsilon_2 + m_2 s_1}{\delta_2 + \varepsilon_2 + m_2} \\
 f_4(s_1, s_2, s_3, s_4) &= \frac{\beta_2 s_3 s_4 + \gamma_2 + \alpha_2 + m_2 s_2}{\beta_2 + \gamma_2 + \alpha_2 + m_2}.
 \end{aligned} \tag{3.5}$$

If the spectral radius of the expectation matrix $\rho(M) > 1$, then the process is supercritical. A formula for the minimal fixed point of (3.5), $(q_1^*, q_2^*, q_3^*, q_4^*)$ can be computed numerically. In the supercritical case, the probability of no major outbreak is approximately

$$\mathbb{P}_0(e_{10}, i_{10}, e_{20}, i_{20}) = (q_1^*)^{e_{10}} (q_2^*)^{i_{10}} (q_3^*)^{e_{20}} (q_4^*)^{i_{20}},$$

where e_{j0} and i_{j0} are the initial number of exposed and infectious individuals in patch j , respectively. The probability of a major outbreak is $1 - \mathbb{P}_0(e_{10}, i_{10}, e_{20}, i_{20})$.

An example with equal dispersal rates for the two populations, $m_1 = m_2$, results in outbreaks in both populations. The probability of an outbreak increases in population 1 but decreases in population 2. In Figure 3.2, population 1 has $\mathcal{R}_{01} = 2.67$ but in population 2 there is lower transmission, higher recovery, and lower mortality, so that $\mathcal{R}_{02} = 0.89$. One infectious individual introduced into population 1 gives a probability for no major outbreak, $\mathbb{P}_0(0, 0, 0, 1) = 0.751$.

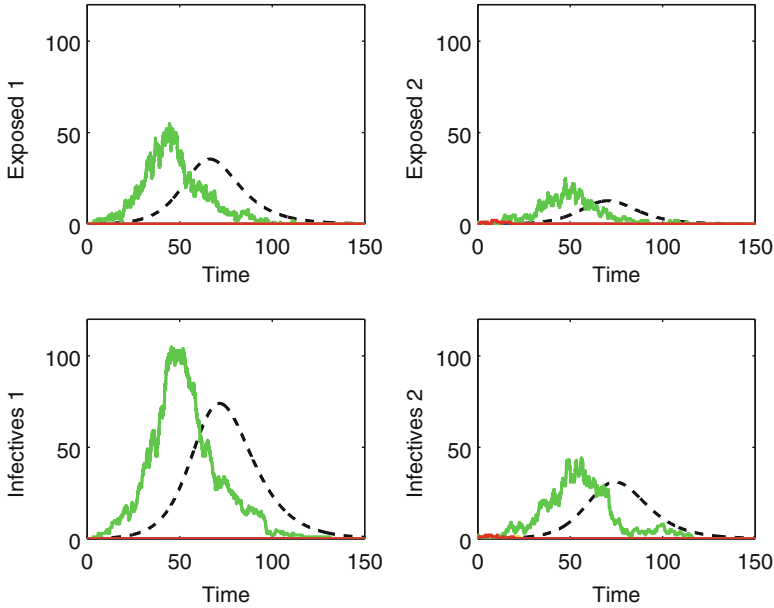


Fig. 3.2 Four sample paths of the SEIR CTMC model with dispersal along with the ODE solution (dashed curve). Parameter values are $\beta_1 = 0.4$, $\beta_2 = 0.2$, $\delta_1 = 0.4 = \delta_2$, $\gamma = 0.1$, $\gamma_2 = 0.05$, $\alpha_1 = 0.05$, $\alpha_2 = 0.025$, $\varepsilon_1 = 0 = \varepsilon_2$, and $m_1 = 0.05 = m_2$. Initial values are $S_1(0) = 500$, $S_2(0) = 499$, $E_1(0) = 0 = E_2(0)$, $I_1(0) = 1$, and $R_1(0) = 0 = R_2(0)$. The basic reproduction numbers for each population are $\mathcal{R}_{10} = 2.67$ and $\mathcal{R}_{02} = 0.89$. Probability of no major outbreak is $\mathbb{P}_0(0, 0, 0, 1) = 0.751$.

With unequal dispersal between the two populations, the probability of an outbreak depends on the direction and magnitude of the dispersal rates (as in the birth-death-dispersal model in Chapter 1). In Figure 3.3, the fixed points $(q_1^*, q_2^*, q_3^*, q_4^*)$ of the pgfs in (3.5) are computed numerically given both m_1 and m_2 lie in the range $[0, 0.2]$. If dispersal is greater toward the population with good healthcare facilities it is possible to eradicate disease in both populations, that is, no major outbreaks occur. Allowing movement out of the poor healthcare region but restricting movement into the region (bottom left graph in Figure 3.3) is a good strategy for disease control, but restricting movement out of the poor healthcare strategy but allowing movement into that region (bottom right graph in Figure 3.3) is a poor strategy.

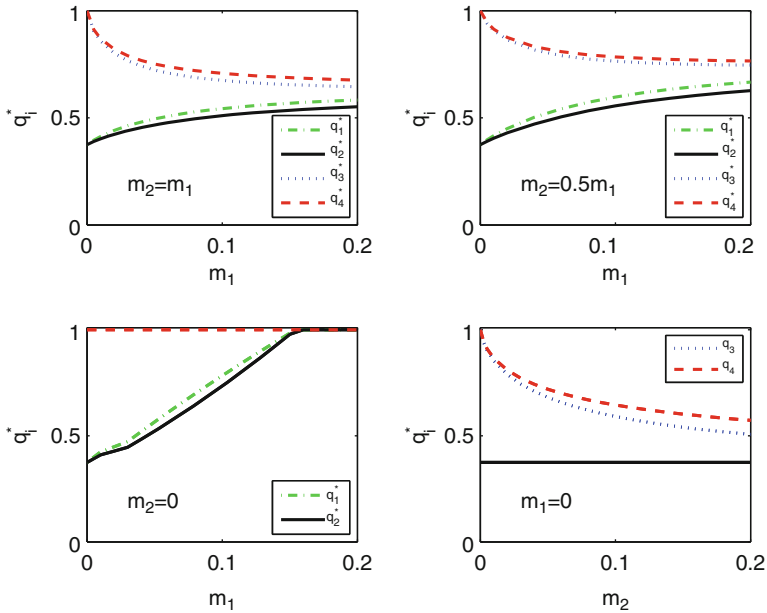


Fig. 3.3 Probability of no major outbreak for a CTMC SEIR epidemic model with dispersal between two populations as estimated by the fixed points q_i^* . Dispersal from population 1 to 2 is m_1 and from population 2 to 1 is m_2 . Parameter values are the same as in Figure 3.2, except for the dispersal parameters: $\beta_1 = 0.4, \beta_2 = 0.2, \delta_1 = 0.4 = \delta_2, \gamma = 0.1, \gamma_2 = 0.05, \alpha_1 = 0.05, \alpha_2 = 0.025, \varepsilon_1 = 0 = \varepsilon_2$.

3.4 Summary

The multi-type branching process application to an SEIR epidemic with dispersal illustrates the importance of controlling movement into and out of particular regions to prevent an outbreak. Although prevention and control measures are more complex in real epidemic or pandemic situations, the basic SIR and SEIR models are often used in conjunction with data to help estimate the potential spread of the disease, e.g., SARS, influenza, and Ebola [11, 22, 33]. The control measures in pandemic situations often include travel restrictions, quarantine, isolation, and drugs such as antiviral medication to prevent infection. Other specific applications of branching processes to infectious disease models include vector-transmitted diseases [4, 6, 9, 18], HIV infection, [12] and bovine respiratory syncytial virus [19].