Branching Processes in Epidemiology

Author:
Timothy Edwin Rice

Supervisors:
Prof. Kostya Borovkov  
Department of Mathematics & Statistics

A/Prof. Robert Day  
Department of Zoology
Contents

1 Introduction 3
   1.1 Toxoplasmosis 3
   1.2 Spanish Influenza 4
   1.3 Abalone Virus Ganglioneuritis 5
   1.4 Crowding and Virulence 5

2 Some Facts from the Theory of Branching Processes 7
   2.1 Continuous-time Multitype Markov Branching Processes 8
       2.1.1 Notation, Definitions and Construction 8
   2.2 Means 11

3 Applications Of Branching Processes To Epidemiology 14
   3.1 Problem Formulation 14
       3.1.1 Modelling Assumptions 14
       3.1.2 Transition Probabilities and their Generating Functions 16
   3.2 Problem Analysis 18
       3.2.1 The Infinitesimal Generator $A$ 18
       3.2.2 The Mean Matrix $M(t)$ 20
       3.2.3 Expected Long-term Behaviour 1: Constant Parameters 21
       3.2.4 Plots and Simulations 1: Constant Parameters 23
       3.2.5 Expected Long-term Behaviour 2: Two-phase Process 27
       3.2.6 Plots and Simulations 2: Two-phase Process 29

4 Conclusion 31

Appendices 33
   A Matlab Code 33
       A.1 simbranch 33
       A.2 means 36
A.3  dsimbranch ........................................ 37
A.4  dmeans .............................................. 41
Abstract

We introduce an open problem in epidemiology and epizootiology, namely, the question of whether increased crowding of a host population causes infectious diseases to become more deadly, and whether deadly infectious diseases, when exposed to less crowded circumstances, will evolve to be more benign. We use the theory of multitype branching processes in continuous-time to develop a model that supports a positive response to this question, and include some simulations that reflect the theoretical results.
Acknowledgements

I would like to thank my Prof Kostya Borovkov for being a source of encouragement, support, and strengthening challenges.

Thanks to A/Prof Robert Day for the fascinating discussions that inspired the present work.

And I can’t say thank you enough to my wife Sasha, without whose gracious and persevering support any of this would have been possible.
Chapter 1

Introduction

“To sustain themselves, they [diseases] need a human population that is sufficiently numerous, and sufficiently densely packed, that a numerous new crop of children is available for infection by the time the disease would otherwise be waning. Hence measles and similar diseases are also known as crowd diseases.”

— From Guns, Germs and Steel by Jared Diamond.

It is a widely-held view – as illustrated by the above quote – that pathogens often face a trade-off in the extent to which they exploit their hosts. Different strains of the same pathogen exploit hosts with greater or lesser attitudes of carpe deum, with more vigorous host subversion leading to more fruitful transmission events, while depleting resources available to the host. A depleted host, rendered immobile and irresponsible by infirmity or death, then presents transmission opportunities less frequently. Clearly, more vigorous pathogens will have fewer opportunities to transmit themselves, but will be better suited to take advantage when such opportunities do arise.

This trade-off is hardly rare or surprising. Indeed, it would be difficult to categorise a biological agent as pathogenic if it did not co-opt host resources for its own benefit. The simplest example concerns cell lysis as the culmination of bacteriophage reproduction (Knox et al., 2001; Dimmock et al., 2007).

1.1 Toxoplasmosis

More interesting, perhaps, than cell lysis by bacteriophages is the disease toxoplasmosis. Montoya and Liesenfeld (2004) specify that domestic felines constitute the
definitive host for the protozoon known as *Toxoplasma gondii*, which is associated with the disease toxoplasmosis. Berdoy et al. indicate that while most mammals are susceptible to toxoplasmosis, cats are the only animals known to shed *T. gondii* oocysts (literally, “egg pouches.”) Cat faeces and infected meat are the two primary vectors for infection... but how do cats become infected in the first place?

Berdoy et al. report behavioural changes in rodents infected with the toxoplasmosis parasite. When uninfected rats encounter cats, their response is as expected for any animal in the presence of its natural predator: strong aversion and extreme caution. This persists in laboratory rats that have not experienced feline predation in several generations.

Rats infected with toxoplasmosis, however, have reduced inhibitions against associating with cats. And that is not all; they in fact prefer or at least seem more interested in areas associated with feline presence. This has furnished the literature with arguments that in order to increase infection-rates in cats, the pathogen evolved the ability to render rats more vulnerable to feline predation. Protozoa which have the most telling impact on rodent behaviour result in the greatest number of new infected felines, which leads to the scattering of more “Trojan” cat faeces, and hence greater numbers of more vulnerable rats.

From the (figurative) point of view of *Toxoplasma gondii*, any other animals who get caught up in the cycle, such as *homo sapiens*, are just collateral damage, if that. The trade-off the pathogen faces is determined more-or-less solely by the interactions of the felines and the rodents. It might not seem like this pathogen in fact faces any such trade-off at all, but of course the spread of *Toxoplasma gondii* to the definitive host depends on the rodent dying. Dead rodents do not reproduce to furnish the pathogen with a new generation of victims. Furthermore, felines demonstrate the greatest interest in lively prey. Therefore, in as much as the protozoa “want” to be endemic or epidemic in their host populations, it is not in their best interest to dull rodents too much.

### 1.2 Spanish Influenza

The biological agent that arguably has the greatest pathogenic impact on humans is influenza. Historical author John M. Barry reports in Barry (2005) that while AIDS had killed perhaps 24 million people by 2005, with another 40 million infected, the Spanish flu alone killed at least 50 million people over a 6 month period. This influenza data is corroborated by Taubenberger and Morens (2006).
It has been suggested (Oxford et al., 2005) that the Spanish flu was partly caused by overcrowding of soldiers during World War I, combined with their subsequent dispersal at the war’s end. It may be noted that the Spanish flu did not hit Australia as hard as it did the rest of the world. This has been attributed to the fact that Australia implemented effective quarantines which prevented incursion of Spanish influenza until 1919 — much later date than it hit the rest of the world.

1.3 Abalone Virus Ganglioneuritis

The belief that overcrowding contributes to so-called “virulence” is indeed prevalent, and there is an intuitive appeal to this belief. According to A/Prof. Robert Day in an oral communication, the hypothetical phenomenon has been observed in Australian abalone stocks with the spread of the Abalone Virus *Ganglioneuritis* (AVG).

This infectious pathogen has been described by Julie Hills (2007), who states that AVG was originally observed in wild Taylor Bay abalone, whence the pathogen has travelled as an invasion wave into virgin stock at the rate of 6-10km per annum. This has already caused millions of dollars of damage to Australia’s abalone industry, which exports the invertebrates as a delicacy, and the rate of new infections is not slowing down.

Julie Hills (2007) specifies that the most likely original source of the infection was a virus already endemic to some wild populations, which were caught and introduced into highly-crowded farms of “naïve stock at the end of 2005. Within months, the farms suffered heavy losses, and in May 2006 – the same month in which Southern Ocean Mariculture needed to completely destock (a loss they have not recovered from) – the first wild cases appeared in the aforementioned Taylor Bay, from which Southern Ocean Mariculture drew its water supply.

1.4 Crowding and Virulence

The chain of cause-and-effect that would implicate Southern Ocean Mariculture and other abalone farmers in the spread of AVG, or the rise of Spanish flu as a result of crowded troop conditions, has proven difficult to put on a rigorous footing. To the best of our knowledge, the current models in the biological literature do not establish a causative relationship that adequately link the combination of overcrowded conditions and subsequent dispersal to the rise of extremely harmful infections.

There do, however exist counterarguments against jumping to the conclusion too
quickly. Namely, it could be surmised that in overcrowded conditions, there is no evolutionary benefit to 
gaining transmissibility if the cost is greater lethality. If a pathogen has many opportunities to spread itself, it does not need 
to make the most of every opportunity presented to it; instead, it may be argued, the infectious agent would be better off spreading itself in a leisurely manner. This question of the effect host crowding has on infectious pathogens is different to concerns about when the trade-off hypothesis is a valid assumption. As a complicated topic with a physical subject, it has been natural to use mathematics to investigate the expected result given that certain assumptions hold.

Numerous attempts have already been made to do this. For example, Anderson and May (1981) state on p461 that “Microparasites with low transmission efficiency will in general only persist within high density populations of hosts,” and “Highly pathogenic parasites will only survive in relatively high-density host populations.” Their dynamical-systems-based models, however, have problems such as deterministic outcomes and assumptions of homogeneous mixing. A/Prof. Day has indicated that while such models have their role, the problem is still open. In the current work, we use the theory of multitype Markov branching processes in continuous-time to attempt to derive some results that have not, so far as we know, occurred elsewhere in the literature.
Chapter 2

Some Facts from the Theory of Branching Processes

Branching Processes model the cardinality and structure of various sets for which there exist generative relationships between elements of the set. Branching processes have a long history of applications in physics and biology; such applications include the number of free neutrons extant in some nuclear reaction, or the extinction probability of family surnames. Both topics were treated in Harris (1989). Graph theory and its generalisations may make rigorous such constructions, and for a good introduction to graph theory, the reader may refer to Chartrand and Oellermann (1993). The present work, however, will not make explicit the graph theoretic formalism, but will instead follow the expositions of Athreya and Ney (1972) and Harris (1989) to construct the processes of interest and present a few useful results.

It should be noted that the association of branching processes with the above-mentioned applications has made conventional certain poetic terms in the branching process literature, regardless of whether or not the author has a specific application in mind. For example, when introducing the Galton-Watson process — originally motivated by questions in genealogy — Athreya and Ney (1972) consistently refer to the basic set elements as “particles.” With this precedent in mind, the present work should also be forgiven for indulging in poetic language when the opportunity presents itself.
2.1 Continuous-time Multitype Markov Branching Processes

Here we focus on Markov Processes on a continuous index set $T = [0, \infty)$ (with elements referred to as “times”), and with a state-space $S$ over $\mathbb{Z}_+^n$, for some specified $n$.

Since there exist at most countably many such vectors, it is possible to establish a one-to-one correspondence between $S$ and $\mathbb{Z}_+^n$. Hence, multitype Markov branching processes are formally equivalent to the single-type jump Markov processes one first encounters in an undergraduate second course in probability.

Despite this formal equivalence, multitype branching processes do warrant separate treatment. Restriction to a single type would not only be counterintuitive for many problems in this area, but would also, in fact, increase the complexity of such problems.

2.1.1 Notation, Definitions and Construction

**Definition 1** ($n$-dimensional continuous-time Markov branching process). If a stochastic process $\{Z(t, \omega) : t \in T\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfies the following properties,

1.) $Z : T \times \Omega \rightarrow \mathbb{Z}_+^n$,

2.) the strong Markov property, so that $Z$ is Markovian with respect to the fields $\mathbb{F}_t := \sigma\{Z(s, \omega) : s \leq t\}$, even when $t = \tau$ for a stopping time $\tau$ (Kiyima, 1997),

3.) the transition probabilities

$$P(i, j; t) := \mathbb{P}\{\omega : Z(s + t, \omega) = j \mid Z(s, \omega) = i\}$$

have generating functions of the form,

$$\sum_{j \in \mathbb{Z}_+^n} P(i, j; t)s^j = \prod_{k=1}^{n} \left[ \sum_{j \in \mathbb{Z}_+^n} P(e_k, j; t)s^j \right]^{i_k},$$

$\forall i = (i_1, i_2, \ldots, i_n) \in \mathbb{Z}_+^n, s \in [0, 1]^n$, and $s, t \geq 0$.

then we call $\{Z(t, \omega) : t \in T\}$ a $n$-dimensional (or multitype with $n$ types) continuous time Markov branching process. Here, $e_k \in \mathbb{Z}_+^n$ has 1 in the $k^{th}$ component and 0 in every other component.
The generating function notation introduces vector-valued powers of vectors, as shown by $s^j$. This is interpreted componentwise, so that $s^j = s^j_1 s^j_2 \ldots s^j_n$, for instance.

Properties 1.) and 2.) define a continuous time Markov process on $\mathbb{Z}_n^+$, while 3.) characterises branching processes with time-homogeneous increments.

**Notation.** For each $t$, $Z(t, \omega)$ is a random variable, so as per the convention for random variables, we will usually not explicitly mention $\omega$ every time, and write $Z(t)$ where $Z(t) \equiv Z(t, \cdot)$.

We write the components of $Z(t)$ unbolded with a subscript indicating the component in question. That is,

$$Z(t) = (Z_1(t), Z_2(t), \ldots, Z_n(t)).$$

If the process starts in some given state, $j$, so that $Z_1(0) = j_1, Z_2(0) = j_2, \ldots, Z_n(0) = j_n$, then we denote the process by $Z^{(j)}(t)$, except in the special case $j = e_i$, when we write instead $Z^{(i)}(t)$ — that is, the superscript is not bolded. In the latter case, we use $P_i$ to denote the associated conditional probability measure; in an analogous manner, $E_i$ will be the expectation conditional on $Z(t) = e_i$.

It shall prove convenient to define $[k] := \{1, 2, \ldots, k\}$ for any $k \in \mathbb{N}$; the inspiration for this notation comes from Wilf (1994). This notation is sometimes used to denote the integer part of $k$, but should the present work require the latter, the notation $\lfloor k \rfloor$ will be used instead.

**Process Construction**

We would like to show that the processes of interest are well-defined and well-behaved in probabilistic terms. Following the exposition in Harris (1989), suppose that a particle of type $i \in [n]$, which exists at some time $t$, has a probability $a_i(t) \epsilon + o(\epsilon)$ of undergoing a transformation during the time interval $(t, t+\epsilon)$. Here, $\epsilon$ is a strictly positive real number but taken as “small” for the purposes of defining the asymptotic term $o(\epsilon)$.

**Assumption 2.1.1 (Exponential Lifespan).** Since we have assumed a homogeneous process, we may write $a_i(t) \equiv a_i$, and recognise that in the limit $\epsilon \to 0$, the transformation probability $a_i(t) \epsilon + o(\epsilon)$ for the interval $(t, t+\epsilon)$ induces an exponentially-distributed lifespan for the particle, with parameter $a_i$. (If, or when, the particle
transforms, it will generate some number in $\mathbb{Z}^n_+$ of new particles. Each new particle must belong to exactly one type from $[n].$)

**Definition 2** (Offspring probability mass function). Following Chapter V in Athreya and Ney (1972), the number of offspring for a type-$i$ parent will follow a discrete probability distribution with masses $p^{(i)}(j_1, j_2, \ldots, j_n) = p^{(i)}(j) := P(Z^{(i)}(\tau) = j)$, where $\tau$ denotes the first split-time of the process, that is, the time when the first particle (which has type $i$) undergoes branching. Note that $\sum_{j \in \mathbb{Z}_+^n} p^{(i)}(j) = 1$. We write the vector containing the probabilities for each particle type as

$$p(j) = (p^{(1)}(j_1, j_2, \ldots, j_n), p^{(2)}(j_1, j_2, \ldots, j_n), \ldots, p^{(n)}(j_1, j_2, \ldots, j_n)).$$ (2.1)

**Notation** (Generating Functions). We will find it convenient to exploit the theory of generating functions, and so assign the notation:

$$g(s) = (g^{(1)}(s), g^{(2)}(s), \ldots, g^{(n)}(s)),$$

where $g^{(i)}(s) = \sum_{j \in \mathbb{Z}_+^n} p^{(i)}(j)s^j$, and,

$$F(i, s; t) = \sum_{j \in \mathbb{Z}_+^n} P(e_i, j; t)s^j.$$

The latter allows us to write,

$$F(s; t) = (F(1, s; t), F(2, s; t), \ldots, F(n, s; t)),$$

where $F(i, s; t) = F(e_i, s; t)$.

Using generating functions, we can now pre-emptively define the “infinitesimal generating functions” (which will simplify the form of the Kolmogorov equations):

$$u(s) = (u^{(1)}(s), u^{(2)}(s), \ldots, u^{(n)}(s)),$$

where $u^{(i)}(s) = a_i[g^{(i)}(s) - s_i].$

Feller (1957) gives a detailed derivation of the Kolmogorov equations for the single-type process, and as stated, any multitype process is equivalent to such a process. Therefore, the forward and backward Kolmogorov equations can be respectively stated here without derivation:

$$\frac{\partial}{\partial t} P(e_i, j; t) = \sum_{k \in [n]} \sum_{j' \in \mathbb{Z}_+^n} k a_k s_k^{-1} p^{(k)}(j) P(e_k, j'; t)s^{j'} - \langle a_k \rangle P(e_i, j; t),$$

$$\frac{\partial}{\partial t} \sum_{j \in \mathbb{Z}_+^n} P(e_i, j; t)s^j = a_i \left[ \sum_{j \in \mathbb{Z}_+^n} p^{(i)}(j) \prod_{k \in [n]} (P(e_k, j; t))^{j_k} - P(e_i, j; t) \right].$$
where \( \langle a_k \rangle = \sum_{k \in [n]} k a_k \).

It should be noted that in the theory of multitype branching processes, it is conventional to rewrite the Kolmogorov equations in terms of \( F(i,s;t) \) by multiplying both sides of the original equations (which have \( P(i,j;t) \) as the dependent variable) by \( s^j \), and summing over \( j \). This leads to the following form for the Kolmogorov equations, reproduced from Athreya and Ney (1972), p201:

\[
\frac{\partial}{\partial t} F(i,s;t) = \sum_{k \in [n]} u^{(k)}(s) \frac{\partial}{\partial s_k} F(i,s;t) \quad \text{— forward equations (2.2)}
\]

\[
\frac{\partial}{\partial t} F(i,s;t) = u^{(i)} F(s,t) \quad \text{— backward equations (2.3)}
\]

for \( i \in [n] \), recalling that,

\[
u^{(i)} [F(s;t)] = a_i [g^{(i)}(F(1,s;t), F(2,s;t), \ldots, F(n,s;t)) - F(i,s;t)].
\]

**Assumption 2.1.2** (Finite Means). It can be shown that a sufficient condition for (2.2) and (2.3) to have unique solutions which diverge unto infinity over finite times only on sets of measure 0 is given by the requirement,

\[
\frac{\partial f^{(i)}(s)}{\partial s_j} \bigg|_{s=(1,1,\ldots,1)} < \infty \quad \forall \, i, j \in [n] \quad (2.4)
\]

The condition given by the set of inequalities (2.4) will be assumed henceforth.

### 2.2 Means

**Definition 3** (Mean Matrix). If we let \( m_{ij}(t) := \mathbb{E}[Z^{(i)}_j(t)] \) — which of course exists by (2.4) — then the mean matrix is the \( n \times n \) matrix with \( m_{ij}(t) \) in row \( i \) and column \( j \). In words, the entry in the \( i \)th row and \( j \)th column of the matrix \( M(t) = (m_{ij}(t))_{i,j \in [n]} \) is the expected number of type-\( j \) individuals alive at time \( t \) given that at time \( t = 0 \), there existed a single type-\( i \) individual.

Using the backward equation (2.3), we can show that the set over \( t \in [0, \infty) \) of all \( M(t) \) is an Abelian semigroup with identity element given by \( \lim_{t \to 0} M(t) = I \) where \( I \) is the \( n \times n \) identity matrix \( (\delta_{ij})_{i,j \in [n]} \).

Since it is trivial to write the identity in the form of a matrix exponential, the semigroup property \( M(s + t) = M(s)M(t) \) then implies the existence of a matrix \( A \)
such that we can write \( M(t) = e^{At} \) for any \( t \in [0, \infty) \). This matrix \( A \) is referred to as the *infinitesimal generator* for the semigroup, and one may infer that

\[
A = A_d \left[ J_g(1) - I \right]
\]

where \( A_d = \text{diag}(\{a_i : i \in [n]\}) \), \( J_g(1) \) is the Jacobian of the vector field \( g(s) \) evaluated at \( s = (1, 1, \ldots, 1) \). That is,

\[
(J_g(1))_{i,j} = \left. \frac{\partial g^{(i)}(s)}{\partial s_j} \right|_{s=(1,1,\ldots,1)}.
\]

**Definition 4** (Positive Regular). If a Markov branching process has a mean matrix \( M(t) \) such that \( \exists t_0 \in (0, \infty) : m_{ij}(t_0) > 0 \quad \forall i, j \in [n] \), then we refer to the process as *positive regular*.

**Assumption 2.2.1** (Irreducibility). In order to exploit the classical Perron-Frobenius theory of positive matrices, Athreya and Ney (1972) impose the assumption of positive regularity. It is easy to see that this will be true for any *irreducible* process, defined (as in Borovkov (2003)) as a process for which every state is accessible from every other state. It will be sufficient for our purposes to assume that \( Z(t) \) is irreducible, and thus irreducibility will be assumed henceforth.

**Notation.** \( \sigma_+ \). Under Assumption 2.2.1, the Perron-Frobenius theory gives us (among other conclusions) that the eigenvalue \( \sigma_+ \) of the infinitesimal generator \( A \) which has the greatest real part will also have 0 imaginary part; \( \sigma_+ \) induces threshold phenomena in the continuous-time multitype Markov branching process in the sense that the latter is considered *supercritical*, *critical* or *subcritical* depending on whether \( \sigma_+ \) is positive, vanishing or negative respectively.

**Definition 5** (Singular). A branching process \( Z(t) \) will be called *singular* if its generating functions define a homogeneous linear system. In a singular process, each time any particle undergoes transformation, it generates exactly one new particle of some sort. Therefore, the total number of particles in the system is constant in time, and so the process is equivalent to a finite jump Markov process. A process that is not singular is called *non-singular*.
Assumption 2.2.2 (Non-singularity). Complementary to irreducibility (assumption 2.2.1), we shall assume $Z(t)$ is non-singular.

Now, $A$ is diagonalisable iff it has an $n$-dimensional eigenspace; in that case, there exists an invertible matrix $P$ such that $P^{-1}AP = D$ where $D = \text{diag}\{\text{eigenvalues of } A\}$ is a diagonal matrix and $P$ is the matrix with eigenvectors of $A$ as its columns. In the case that $A$ is diagonalisable, we can calculate $M(t)$ using,

$$M(t) = e^{tA}$$
$$= \sum_{i=0}^{\infty} (tA)^i$$
$$= \sum_{i=0}^{\infty} t^i [PDP^{-1}]^i$$
$$= \sum_{i=0}^{\infty} t^i \left[ PDP^{-1} PDP^{-1} \cdots PDP^{-1} \right]_{i \text{ times}}$$
$$= \sum_{i=0}^{\infty} t^i PD^i P^{-1}$$
$$= P \left[ \sum_{i=0}^{\infty} t^i D^i \right] P^{-1}$$
$$= P \text{diag} \left[ \sum_{i=0}^{\infty} t^i \sigma_j^i \right] P^{-1}$$

where $\sigma_j^i$ is the $jth$ eigenvalue of $A$, raised to the $ith$ power
$$= P \text{diag}(e^{\sigma_j t}) P^{-1}$$

(2.6)
Chapter 3

Applications Of Branching Processes To Epidemiology

3.1 Problem Formulation

We wish to use the continuous-time multitype Markov branching processes formalism to address the following:

**Hypothesis.** As a host population becomes more crowded, it will support more lethal pathogens (over benign alternatives). When crowding decreases, the benign strains will again become dominant.

**A Note On Language.** It should be noted that in the epidemic literature, the term *virulence* often sees print. When used as the name of a model parameter, it may refer to either the lethality of a pathogen, or its transmissibility. This leads to a conflation of the two concepts. For example, see Table 1 on page 1443 of Herre (1993), where Dr. Herre says “More virulent nematode species are associated with greater opportunities for transmission.”

Herein, the terms lethality and transmissibility will be used as names for model parameters. The term *virulent* may refer to a pathogen which combines high lethality with high transmissibility, when compared to some other (“benign”) pathogen.

3.1.1 Modelling Assumptions

**Assumption 3.1.1 (Number of Types).** Since we wish to clearly establish any effect due to contrasting lethality and transmissibility, it would be unwise, within the theory of multitype branching processes in continuous-time, to extend our attention to the
problem with more than two types. Since there are two scales, the problem with four
types — low lethality and low transmissibility, low lethality and high transmissibility,
high lethality and low transmissibility, and high scores on both scales — would also
be of interest. However, this approach seemed inadvisable for several reasons:

1. We can modify the parameters of the simpler model to compare any two at
   a time of the four combinations described above — and this will more clearly
   emphasise the differences in behaviour (if any) for each pair;

2. The principle of parsimony favours simpler models — this is no mere philosoph-
   ical platitude, but a foundation of successful mathematical modelling;

3. It seems reasonable to predict that the four-type model will not generate new
   insights — it employs exactly the same mathematical formalism.

4. Supposing that there were more than two pathogenic strains, it would be mathe-
   matically trivial to quotient the set of all strains into two partitions. Infections
   belonging to one partition or the other would be called virus 1 or virus 2 as
   appropriate.

Assumption 3.1.2 (Number of Simultaneous New Infections). We wish the model
to have just a few parameters, in order to more clearly emphasise the difference made
by the important factors. Therefore, we assume that each infected individual can
transmit the pathogen to at most one new host at a time. Modelling a pathogen
with superior transmissibility shall involve increasing the probability of causing new
infections, not increasing the number of possible infections.

This assumption is justified from a physical perspective because in fact it seems
unreasonable to suppose that a single individual could cause two or more new in-
fections simultaneously. The model does allow an established host to go on creating
new infections, and so it is possible for a host to establish two or more infections over
arbitrarily-short time intervals.

One might note that this assumption leads to a two-type analogue to the classical

Assumption 3.1.3 (Size and Behaviour of Population). Epidemic models normally
presuppose the existence of a set of “susceptible” individuals. See, for example,
Anderson and May (1981). These susceptibles are supposed to interact with the
infected population, and the literature usually makes assumptions (either implicit or
explicit) about how this takes place. For example, to justify models in the form of
a smooth dynamical system, as used in Anderson and May (1981), one must assume that populations are subject to “homogeneous mixing,” as the equations draw on the notion of a law of mass action, borrowed from the theory of chemical reactions.

For the continuous-time multitype Markov branching process model, we shall assume the population of susceptibles is countably infinite and that infection patterns are defined exactly by the equations which describe the branching process model. The physical mechanism of infection, as well as the activities of the susceptible population, are both assumed not so much irrelevant as captured by the stochasticity of the branching process. It is, however, necessary to have a parameter that will control crowding in the total population. This will be introduced below.

This assumption, that the population is effectively infinite, may prove appropriate when modelling epidemics in populations for which the number of susceptibles is quite large, and in any case considerably greater than the number of infected individuals. It should be noted that where the model predicts the rise of a particular deadly and infectious pathogen, then the model will only be valid over very short time periods, since such a “virulent” specimen will affect the real system in such a way that the branching process model is no longer valid.

3.1.2 Transition Probabilities and their Generating Functions

Suppose that each infected host has an exponentially-distributed lifespan, as well as an exponentially-distributed time until it transmits the pathogen to a new host. It can only do the latter if it doesn’t die first. Hence the time before a given host induces a state-change in the system is the minimum of two independent exponential random variables.

A basic result holds that for any independent exponentially-distributed random variables, \( \tau_1 \) and \( \tau_2 \) with intensity parameters \( \lambda_1 \) and \( \lambda_2 \) respectively, their minimum is another exponential random variable, \( \tau \), with intensity parameter given by \( \lambda = \lambda_1 + \lambda_2 \). A branching process changes states at the end of this time, according to a probability distribution derived from conditioning on a random variable which indicates whether the state-change occurred due to host-removal or pathogen-transmission.

**Notation.** \( \alpha, \beta, \mu \) We shall consider the case where there exist two pathogenic varieties, *virus 1* and *virus 2*. Each pathogen has associated with it a lifespan intensity parameter \( \alpha_i \), a transmission probability \( \beta_i \), and a mutation probability \( \mu_i \), \( i \in \{1, 2\} \).
To clearly distinguish between the two pathogens, we shall rewrite $\beta_1 = \beta$ and $\beta_2 = \gamma$.

**Notation.** $\lambda$ It remains necessary to introduce a final parameter, to determine crowding in the total population. Although the (countably infinite) susceptible population never diminishes in size, infected hosts might be thought of as having more frequent or less frequent opportunities to pass on their pathogens. This will be reflected in the *meeting intensity* parameter $\lambda$, where $\lambda > 0$.

**Assumption 3.1.4** (Time scaling). It should be noted, and will henceforth be assumed, that (without loss of generality) we can change the units of time and assume $\alpha_1 = 1$; we can then drop the subscript on the $\alpha_2$ to write simply $\alpha$, with $\alpha > 0$.

**Assumption 3.1.5** (Non-trivial Mutation Probabilities). The irreducibility assumption imposes the constraint that $\mu_1, \mu_2 \in (0, 1)$. The physical interpretation of this is that the two pathogens are closely related, perhaps in the same sense that two strains of influenza are related, and relatively benign strains may spontaneously give rise to more virulent strains, or vice-versa.

**Assumption 3.1.6** (Identical Mutation Probabilities). To keep the model as simple as possible, and to prevent the question of different mutation rates from bringing the results into question, we set $\mu_1 = \mu_2 = \mu$ for some $\mu \in (0, 1)$.

**Notation.**

Now, in the two-type branching process we have $p(j) = (p^{(1)}(j_1, j_2), p^{(2)}(j_1, j_2))$, 17
and hence the above assumptions imply that, for our model,

\[
\begin{align*}
\mathbf{p}(0, 0) &= \left( \frac{1}{\lambda+1}, \frac{\alpha}{\lambda+\alpha} \right), \\
\mathbf{p}(1, 0) &= \left( \frac{(1-\beta)\lambda}{\lambda+1}, 0 \right), \\
\mathbf{p}(2, 0) &= \left( \frac{(1-\mu)\beta\lambda}{\lambda+1}, 0 \right), \\
\mathbf{p}(1, 1) &= \left( \frac{\mu\beta\lambda}{\lambda+1}, \frac{\mu\gamma\lambda}{\lambda+\alpha} \right), \\
\mathbf{p}(0, 2) &= \left( 0, \frac{(1-\mu)\gamma\lambda}{\lambda+\alpha} \right), \\
\mathbf{p}(0, 1) &= \left( 0, \frac{(1-\gamma)\lambda}{\lambda+\alpha} \right). 
\end{align*}
\]  

(3.1)

For instance, the probability that a host infected with virus 1 “transforms” into two hosts infected with virus 1 is \(\frac{(1-\mu)\beta\lambda}{\lambda+1}\), as shown by the first component of the vector \(\mathbf{p}(2, 0)\). The second component of that same vector is 0, which reflects the fact that we do not allow a host with virus 2 to instantaneously become two hosts with virus 1. (The latter could happen over an arbitrarily short time period by combining a sequence of mutation and infection events.)

From the above vectors of probabilities, we obtain the vector of generating functions,

\[
g(s) = \left( \frac{1 + (1-\beta)\lambda s_1 + (1-\mu)\beta\lambda s_1^2 + \mu\beta\lambda s_1 s_2}{\lambda+1}, \frac{\alpha + \mu\gamma\lambda s_1 s_2 + (1-\mu)\gamma\lambda s_2^2 + (1-\gamma)\lambda s_2}{\lambda+\alpha} \right).
\]  

(3.2)

### 3.2 Problem Analysis

#### 3.2.1 The Infinitesimal Generator \(A\)

If we treat \(g(s)\) as a vector field on the set of \(s\) in the unit square \([0,1]^2\), we can calculate its Jacobian matrix,

\[
J_g(s) = \begin{bmatrix}
\frac{(1-\beta)\lambda+2(1-\mu)\beta\lambda s_1 + \mu\beta\lambda s_2}{\lambda+1} & \frac{\mu\beta\lambda s_1}{\lambda+1} \\
\frac{\mu\gamma\lambda s_2}{\lambda+\alpha} & \frac{\mu\gamma\lambda s_1 + 2(1-\mu)\gamma\lambda s_2 + (1-\gamma)\lambda}{\lambda+\alpha}
\end{bmatrix},
\]  

(3.3)
which leads to

\[
J_g(1) = \begin{bmatrix}
\frac{(1-\beta)\lambda + 2(1-\mu)\beta \lambda s_1 + \mu \beta \lambda s_2}{\lambda + 1} & \frac{\mu \beta \lambda s_1}{\lambda + 1} \\
\frac{\mu \gamma \lambda s_2}{\lambda + \alpha} & \frac{\mu \gamma \lambda s_1 + 2(1-\mu)\gamma \lambda s_2 + (1-\gamma)\lambda}{\lambda + \alpha}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\frac{(1+\beta)\lambda - \mu \beta \lambda}{\lambda + 1} & \frac{\mu \beta \lambda}{\lambda + 1} \\
\frac{\mu \gamma \lambda}{\lambda + \alpha} & \frac{(1+\gamma)\lambda - \mu \gamma \lambda}{\lambda + \alpha}
\end{bmatrix}.
\]  (3.4)

The \(ij^{th}\) entry of this certainly equates to \(\frac{\partial g(1)}{\partial s_j}\), \(i, j \in \{1, 2\}\); this of course is the expected number of type-\(j\) infections produced during the transformation of a host infected with virus \(i\). Therefore,

\[
B := J_g(1) - I
\]

\[
= \begin{bmatrix}
\frac{(1+\beta)\lambda - \mu \beta \lambda}{\lambda + 1} - 1 & \frac{\mu \beta \lambda}{\lambda + 1} \\
\frac{\mu \gamma \lambda}{\lambda + \alpha} & \frac{(1+\gamma)\lambda - \mu \gamma \lambda}{\lambda + \alpha} - 1
\end{bmatrix},
\]  (3.5)

and

\[
A = A_d B
\]

\[
= \begin{bmatrix}
\lambda + 1 & 0 \\
0 & \lambda + \alpha
\end{bmatrix} \begin{bmatrix}
\frac{(1+\beta)\lambda - \mu \beta \lambda}{\lambda + 1} - 1 & \frac{\mu \beta \lambda}{\lambda + 1} \\
\frac{\mu \gamma \lambda}{\lambda + \alpha} & \frac{(1+\gamma)\lambda - \mu \gamma \lambda}{\lambda + \alpha} - 1
\end{bmatrix}
\]

\[
= \begin{bmatrix}
(1-\mu)\beta \lambda - 1 & \mu \beta \lambda \\
\mu \gamma \lambda & (1-\mu)\gamma \lambda - \alpha
\end{bmatrix}.
\]  (3.6)

The matrix \(A_d\) was defined in Section 2.2 as containing on its diagonal the intensity parameters for the split-times induced by each pathogen type. In the present case, these parameters correspond to the intensity for the minima of two independent exponential random variables, one with intensity \(\lambda\), while the other has intensity 1 (for virus 1) or \(\alpha\) (for virus 2.) Hence \(A_d(1, 1) = \lambda + 1\) and \(A_d(2, 2) = \lambda + \alpha\).

One can easily find that \(A\) has eigenvalues \(\sigma_+\) and \(\sigma_-\) given by,

\[
\sigma_{\pm} = \frac{1}{2} \left[ (1-\mu)(\gamma + \beta)\lambda - 1 - \alpha \pm \Delta \right],
\]  (3.7)

where

\[
\Delta = \sqrt{[(\gamma - \beta)(1-\mu)\lambda + 1 - \alpha]^2 + 4\gamma \beta \mu^2 \lambda^2}.
\]

Likewise, \(A\) has eigenvectors \(\mathbf{u}\) and \(\mathbf{v}\), with respective correspondence to \(\sigma_+\) and
\( \sigma_\pm \), given by,

\[
\begin{align*}
\mathbf{u} &= \begin{bmatrix}
\frac{2\mu\beta\lambda}{(1-\mu)(\gamma-\beta)(\lambda+1-\alpha+\Delta)} \\
1 \\
\frac{2\mu\beta\lambda}{(1-\mu)(\gamma-\beta)(\lambda+1-\alpha-\Delta)} \\
1
\end{bmatrix}, \\
\mathbf{v} &= \begin{bmatrix}
\frac{2\mu\beta\lambda}{(1-\mu)(\gamma-\beta)(\lambda+1-\alpha+\Delta)} \\
1 \\
\frac{2\mu\beta\lambda}{(1-\mu)(\gamma-\beta)(\lambda+1-\alpha-\Delta)} \\
1
\end{bmatrix}
\end{align*}
\]  
(3.8) 
(3.9)

### 3.2.2 The Mean Matrix \( M(t) \)

Recall from Chapter 2 the definition of the mean matrix \( M(t) \). It was defined by assigning to its \((i, j)^{\text{th}}\) entry the expected number of type-\( j \) individuals alive at time \( t \), given that there was a single type-\( i \) individual at time \( t = 0 \).

Since \( \beta, \gamma, \mu \in (0, 1) \) and \( \alpha, \lambda > 0 \), we have that \( \Delta \in \mathbb{R}_+ \), which implies the nice property that \( \mathbf{u} \) and \( \mathbf{v} \) are linearly independent. (Of course, this would follow from the fact that the eigenvalues are distinct, but the fact is more clearly demonstrated by comparing the eigenvectors themselves.) Hence, the dimension of the eigenspace of \( A \) is equal to the number of types in the process, which allows us to diagonalise \( A \) and write \( M(t) \) explicitly,

\[
M = P \begin{bmatrix} e^{\sigma_+ t} & 0 \\ 0 & e^{\sigma_- t} \end{bmatrix} P^{-1},
\]  
(3.10)

where

\[
P = \begin{bmatrix} \mathbf{u} & \mathbf{v} \end{bmatrix}
\]  
(3.11)

contains the eigenvectors of \( A \) (as columns). Since the second component of both eigenvectors is 1, we can write the matrix \( P \) as:

\[
P = \begin{bmatrix} u & v \\ 1 & 1 \end{bmatrix},
\]  
(3.12)

where \( u \) and \( v \) (note: unbolded) are the first components of the eigenvectors \( \mathbf{u} \) and \( \mathbf{v} \), respectively. Of course, from this we can write,

\[
P^{-1} = \frac{1}{u-v} \begin{bmatrix} 1 & -v \\ -1 & u \end{bmatrix},
\]  
(3.13)
and hence

$$M = \frac{1}{u-v} \begin{bmatrix} u & v \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{\sigma t} & 0 \\ 0 & e^{-\sigma t} \end{bmatrix} \begin{bmatrix} 1 & -v \\ -1 & u \end{bmatrix} = \frac{1}{u-v} \begin{bmatrix} u e^{\sigma t} - v e^{-\sigma t} & u v e^{\sigma t} - uv e^{\sigma t} \\ u e^{\sigma t} - v e^{-\sigma t} & u v e^{\sigma t} - uv e^{\sigma t} \end{bmatrix}. \quad (3.14)$$

It is easy to see from the diagonalisation of $M(t)$ that the process can be regarded (or defined) as supercritical, critical, or subcritical, depending on whether $\sigma_+$ is greater than, equal to, or less than, 0. It might be noted that $\sigma_+$ is monotonic increasing in $\lambda$ at an asymptotically linear rate, and we can make $\lambda$ arbitrarily large. Therefore, we can state the first prediction of the model: that by sufficiently increasing the crowding of a population, we can cause a pathogenic pair, virus 1 and virus 2, to switch from subcriticality to supercriticality. This is hardly different to the single-type result, but it is heartening that our multitype model reproduces anticipated behaviour.

### 3.2.3 Expected Long-term Behaviour 1: Constant Parameters

The quantity $R(t)$ defined by

$$R(t) := \frac{M_{12}(t)}{M_{11}(t)}, \quad (3.15)$$

gives the ratio at time $t$ of the expected number of individuals infected with virus 2 to that infected with virus 1, given that the initial infected population consisted of a single individual infected with virus 1. We can therefore write

$$R(t) = \frac{u v e^{\sigma t} - uv e^{\sigma t}}{u e^{\sigma t} - v e^{-\sigma t}} = \frac{u v e^{\sigma t} - uv e^{\sigma t}}{u e^{\sigma t} - v e^{-\sigma t}} = \frac{u v e^{\sigma t}}{u e^{\sigma t}}.$$

$$\Rightarrow \lim_{t \to \infty} R(t) = -\frac{uv}{u} = -v. \quad (3.16)$$
Since a consideration of the denominator of \( v \) shows that \( v \) is always negative, we can define the ratio of the expected long-time number of individuals infected by virus 2 to that of individuals infected by virus 1 as,

\[
R := \lim_{t \to \infty} R(t)
\]

\[
= |v|
\]

\[
= \frac{2\mu\beta\lambda}{\Delta - (1 - \mu)(\gamma - \beta)\lambda - 1 + \alpha}
\]  

(3.17)

There are several important consequences of having such an explicit form for the ratio. We can plot \( R \) as a function of the parameters \((\alpha, \beta, \gamma, \lambda, \mu)\) and compare these plots to simulation results, and we can calculate the parametric critical points that will cause the ratio to switch from subunity to superunity.

The next set of results will be obtained from solving \( R = 1 \) for the various parameters. This will yield thresholds that quotient scenarios on cases where virus 2 is dominant, and where it is not. Firstly, for \( \alpha \):

\[
R = 1
\]

\[
\Rightarrow \frac{2\mu\beta\lambda}{\Delta - (1 - \mu)(\gamma - \beta)\lambda - 1 + \alpha} = 1
\]

(3.18)

\[
\Rightarrow 2\mu\beta\lambda = \Delta - (1 - \mu)(\gamma - \beta)\lambda + 1 - \alpha
\]

\[
\Rightarrow (\gamma - \beta)(1 - \mu)\lambda + 1 - \alpha + 2\mu\beta\lambda = \Delta
\]

From this, we can obtain,

\[
[(\gamma - \beta)(1 - \mu)\lambda + 1 - \alpha]^2 + 4\gamma\beta\mu^2\lambda^2 = [(\gamma - \beta)(1 - \mu)\lambda + 1 - \alpha]^2
\]

\[
+ 4\beta\mu\lambda[(\gamma - \beta)(1 - \mu)\lambda + 1 - \alpha]
\]

\[
+ 4\beta^2\mu^2\lambda^2
\]

\[
\Rightarrow 4\mu\beta\lambda\alpha = 4\beta\mu\lambda[(\gamma - \beta)(1 - \mu)\lambda + 1]
\]

\[
+ 4\beta^2\mu^2\lambda^2 - 4\gamma\beta\mu^2\lambda^2
\]

\[
\Rightarrow \alpha = (\gamma - \beta)(1 - \mu)\lambda + 1 - (\gamma - \beta)\mu\lambda
\]

\[
= (\gamma - \beta)(1 - 2\mu)\lambda + 1.
\]

(3.19)

Since the left-hand side of equation (3.18) is decreasing in \( \alpha \), the threshold in (3.19) must provide an upper limit on how intently virus 2 kills its victims, if it is to be dominant. That is, to have \( R > 1 \), we require,

\[
\alpha < (\gamma - \beta)(1 - 2\mu)\lambda + 1.
\]

(3.20)
Note that when $\gamma - \beta$ is positive and $\mu < 0.5$, the cap is monotonic increasing in $\lambda$. This supports the contention that overcrowding favours more deadly pathogens: when virus 2 has better transmission probabilities than virus 1, virus 2 can afford to be more deadly, and it can afford to be more and more deadly as overcrowding increases.

Likewise, considering (3.20) allows us to formulate a threshold for $\lambda$ as,

$$0 < (\gamma - \beta)(1 - 2\mu)\lambda + 1 - \alpha$$

$$\Rightarrow \quad \lambda > \frac{\alpha - 1}{(\gamma - \beta)(1 - 2\mu)}$$

(3.21)

This indicates that not only can more deadly pathogens proliferate as overcrowding increases, but overcrowding will cause more deadly pathogens to “switch on.” Observing inequalities (3.20) and (3.21) when $\gamma = \beta$ (or in the limit $\gamma \to \beta$ for (3.21)) shows that if the pathogens have the same transmission probabilities, it will never favour virus 2 to have $\alpha \geq 1$. This is an important point — it shows that we could not have come up with a simpler model but just as effective model by assuming $\gamma = \beta$. It shows that the model we do have has the least number of parameters required to capture the phenomenon of interest.

Finally, we now formulate a threshold for the “probability excess” $\gamma - \beta$:

$$0 < (\gamma - \beta)(1 - 2\mu)\lambda + 1 - \alpha$$

$$\Rightarrow \quad \gamma - \beta > \frac{\alpha - 1}{(1 - 2\mu)\lambda}$$

(3.22)

The right-hand side of the inequality (3.22) is decreasing in $\lambda$. The implication of this is that no matter how deadly virus 2 is (that is, no matter how great the value of $\alpha$), we can always find a $\lambda$ such that the probability difference $\gamma - \beta$ need only be arbitrarily small in order to ensure that virus 2 proliferates in greater numbers than virus 1. In practical terms, if a population is sufficiently crowded, the model predicts that more deadly pathogens do not need to exhibit measurably better transmission probabilities in order to dominate.

### 3.2.4 Plots and Simulations 1: Constant Parameters

We now consider some simulation results and compare with the analytic results obtained above. These plots were all obtained using Matlab r2007b. To see the relevant Matlab code, refer to Appendix A. It should be noted that in most simulations, the
pathogens die out very quickly. Where four plots appear in the same image, the two on the left-hand side give the simulation results, the two on the right-hand side represent the theoretical predictions, the two topmost images represent absolute numbers of individual infected with each pathogen, and the two lower figures correspond to the ratio of virus 2 to virus 1.

Figure 3.1 gives an example of the typical simulation, in which both pathogens die out before they can become established. Figure 3.2 shows that for these parameters, we can usually expect the pathogens to become supercritical with virus 2 dominant.

It may be observed by comparing the left-hand side to the right-hand side in Figures 3.1, 3.4 and 3.6 that the simulations do compare reasonably well with the predicted results. Their trajectories do seem a little higher than the corresponding expected values, but it needs to be borne in mind that 3.4 and 3.6 were each selected from several candidate simulations based on the criterion of non-termination of the process. They might be thought of as realisations from the related process in which $Z(t)$ is conditioned on the event $Z(t) > 0$.

The observant reader will further note that the simulations are of fairly short duration. This is because as $|Z(t)|$ grows, successive split times are forced closer and
Figure 3.2: Expected long-term behaviour for Figure 3.1.

Figure 3.3: Expect supercritical with virus 1 dominant
Figure 3.4: Simulation with parameters from Figure 3.3

Figure 3.5: Expect supercritical with virus 2 dominant
closer together. Each split time requires a new computation. Therefore, the simulations yield less and less output per second of computation time once $|Z(t)|$ settles into exponential growth. This trade-off means it is not justified to run simulations for excessive lengths of time; once they have provided some qualitative insight, their work is done.

### 3.2.5 Expected Long-term Behaviour 2: Two-phase Process

We have shown in subsections 3.2.3, 3.2.4 that crowding not only favours so-called supercritical processes, but it furthermore favours more deadly pathogens. That model, however, relied on parameters that did not change over the course of the process. For our application of interest, we wish to consider what effect a sudden change in parameters will have. It would be suggestive if we could show that making a population more crowded really will cause more deadly infectious diseases to suddenly switch on.

As our process makes the admittedly hefty assumptions of branch independence and exponentially-distributed lifetimes, it is not too hard to condition the process on
\( \mathbf{Z}(t_C) = (V_1, V_2) \) for some time \( t = t_C \). Then for \( t \geq t_C \), we can use the semigroup property to write,

\[
M^C(t) = \begin{bmatrix} V_1 & V_2 \\ V_1 & V_2 \end{bmatrix} e^{A(t-t_C)}
\]

\[= \frac{1}{u-v} \begin{bmatrix} V_1 & V_2 \\ V_1 & V_2 \end{bmatrix} \begin{bmatrix} u & v \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{\sigma_+(t-t_C)} & 0 \\ 0 & e^{\sigma_-(t-t_C)} \end{bmatrix} \begin{bmatrix} 1 & -v \\ -1 & u \end{bmatrix}
\]

\[= \frac{1}{u-v} \begin{bmatrix} V_1 & V_2 \\ V_1 & V_2 \end{bmatrix} \begin{bmatrix} u e^{\sigma_+(t-t_C)} - v e^{\sigma_-(t-t_C)} & u v e^{\sigma_-(t-t_C)} - u v e^{\sigma_+(t-t_C)} \\ e^{\sigma_+(t-t_C)} - e^{\sigma_-(t-t_C)} & u e^{\sigma_-(t-t_C)} - u e^{\sigma_+(t-t_C)} \end{bmatrix}
\]

\[= \frac{1}{u-v} \begin{bmatrix} (uV_1 + V_2) e^{\sigma_+(t-t_C)} - (vV_1 + V_2) e^{\sigma_-(t-t_C)} \\ (uV_1 + V_2) e^{\sigma_+(t-t_C)} - (vV_1 + V_2) e^{\sigma_-(t-t_C)} \\ u(vV_1 + V_2) e^{\sigma_-(t-t_C)} - v(uV_1 + V_2) e^{\sigma_+(t-t_C)} \\ u(vV_1 + V_2) e^{\sigma_-(t-t_C)} - v(uV_1 + V_2) e^{\sigma_+(t-t_C)} \end{bmatrix}
\]

(3.23)

It may be noted that the second row in \( M^C(t) \) is the same as the first. This is only natural in a Markov process. Having conditioned on the values of \( \mathbf{Z}(t) \) at some \( t > 0 \), the process “forgets” its past. Hence we can define the conditional ratio of expectations, analogous to the \( R(t) \) from subsection 3.2.3 at time \( t \geq t_C \) as

\[
R_C(t) = \frac{M_{12}^C(t)}{M_{11}^C(t)} = \frac{M_{22}^C(t)}{M_{21}^C(t)}
\]

\[= \frac{u(vV_1 + V_2) e^{\sigma_-(t-t_C)} - v(uV_1 + V_2) e^{\sigma_+(t-t_C)}}{(uV_1 + V_2) e^{\sigma_+(t-t_C)} - (vV_1 + V_2) e^{\sigma_-(t-t_C)}}
\]

\[= \frac{e^{\sigma_+(t-t_C)}}{e^{\sigma_+(t-t_C)}} \frac{u(vV_1 + V_2) e^{(\sigma_--\sigma_+)(t-t_C)} - v(uV_1 + V_2)}{(uV_1 + V_2) e^{(\sigma_--\sigma_+)(t-t_C)} - (vV_1 + V_2) e^{(\sigma_--\sigma_+)(t-t_C)}}
\]

\[= \frac{u(vV_1 + V_2) e^{(\sigma_--\sigma_+)(t-t_C)} - v(uV_1 + V_2)}{uV_1 + V_2 - (vV_1 + V_2) e^{(\sigma_--\sigma_+)(t-t_C)}}
\]

\[\Rightarrow \lim_{t \to -\infty} R_C(t) = -\frac{-v(uV_1 + V_2)}{uV_1 + V_2} = -v = |v|
\]

(3.24)

This is exactly the same limit as for the unconditioned result. By the Markov property, we can now easily suppose that prior to time \( t = t_C \), the Markov branching process was operating under some different set of parameters. Thus, if prior to time \( t_C \), the population crowding ensured the domination of benign strains such as virus 1,
but then at time $t_C$ the rate of susceptible contacts increased, then in the long run, we can still expect the more deadly virus 2 to dominate. Conversely, if environmental conditions are favouring virus 2, the model shows that we can expect to restore the domination of benign strains. Furthermore, by independence of the branches, if the pathogenic system is supercritical in the regime prior to $t_C$, then it may be made subcritical for $t \geq t_C$, or vice-versa (provided $||V_1, V_2|| \geq 0$, of course.)

### 3.2.6 Plots and Simulations 2: Two-phase Process

We can see in Figures 3.7, 3.8 and 3.9 that simulations for the two-phases process also accord well with the theory.
Figure 3.8: Ratio of expectations based on an initial population with 20 type-1 individuals and 0 type-2 individuals.

Figure 3.9: Simulation with parameters from Figure 3.8.
Chapter 4

Conclusion

Biologists have provided anecdotal evidence that overcrowded conditions cause the appearance of more deadly strains of infectious diseases. How this phenomenon might be modelled remains an open question. We sought the simplest model that would capture the observed behaviours, with the most transparent interpretation of the parameters. Simpler models are less prone to errors (whether logical or statistical,) and they are more tractable from the point of view of finding numerical or analytical solutions.

We used the theory of multitype Markov branching processes in continuous-time to model the competition between two infectious pathogens when subject to different circumstances. At least two viruses were required to reproduce the phenomenon, but using more than two would have obscured the results.

The model we developed captured the phenomenon in question when a minimum of five dimensionless parameters were used:

1.) $\alpha$ to compare the removal rate of virus 2 to that of virus 1;
2.) $\beta$ to describe the transmissibility of virus 1 (as a probability per opportunity for a new infection)
3.) $\gamma$ to describe the transmissibility of virus 2, in terms similar to $\beta$ for virus 1;
4.) $\lambda$ to describe the susceptibility of the host population;
5.) $\mu$ to model mutation probabilities, ensuring that the model is non-trivial.

The model also had the nice property that its associated matrices are always diagonalisable. Diagonalisable matrices greatly simplify the task of computing a matrix exponential, and hence it was possible to construct the matrix of expected
numbers of individuals infected with each virus type. This simplified the task of comparing simulations with theory.

Using the Markov property, we constructed predictions for the effect of changing parameters. The model predicts that increasing the crowding of a population can cause more deadly infections to suddenly “switch on.” Conversely, if a population is in the grip of a deadly infection, the model predicts that if it is possible to control the population density, then it may be possible to restore the dominance of more benign strains. An alternate prediction is that for a supercritical process dominated by virus 2, the susceptible population will be quickly thinned out, so that in such a situation, we can expect that the emergence of deadly pandemics will only ever be short-lived, as they lead to alterations in the parameter $\lambda$. 
Chapter Appendices

A Matlab Code

The following two Matlab programs contain commenting to the right of any "%" character. Both programs commence with a commented description of their purpose, and commenting throughout should help illuminate their inner structure.

A.1 simbranch

function A = simbranch(r,b,g,m,l)

% Simulates a 2-type cts-time branching process interpreted as a
% competition between viruses in a host population.
%
% Assumes that the initial population has one host infected with virus 1.
%
% Input values:
% r: ratio of removal rates (due to death or recovery) bw viruses 1 & 2
% (r>1 => virus 2 is associated with a higher removal rate.)
% b: probability that an encounter with virus 1 -> infection
% g: probability that an encounter with virus 2 -> infection
% m: probability of mutation in a new infection
% l: exposure rate (due to population density, for instance)

N = [1; 0];  % N will track the number of each type
T = [0];    % T will track the split times
P = [1, b; r, g];  % P stores the pathogen parameters:
removal rate, infection probability

\[ R = \frac{N(2,1)}{N(1,1)}; \]

\[ M = \expm\left(\begin{bmatrix} (1-m)b*l-1 & m*b*l \\ m*g*l & (1-m)*g*l-r \end{bmatrix}\right)T(1); \]

\[ O = \begin{bmatrix} M(1,1); M(1,2) \end{bmatrix}; \]

\[ s = 0.8; \]

fs = get(0,'ScreenSize'); \% fs = "fullscreen"

fsm = [(1-s)*fs(3)/2 (1-s)*fs(4)/2 s*fs(3) s*fs(4)];

figure('Position',fsm,'Name');
title('Simulation vs. Prediction based on 1 initial host with pathogen type 1');

while(any(N(:,1)))

\% The following if/else statements induce the branchings...

D = N(:,1);

check = all(N(:,1));

if(all(N(:,1)))

\[ t_1 = \frac{-\log(rand)}{(N(1,1)\cdot (l+1))}; \]
\[ t_2 = \frac{-\log(rand)}{(N(2,1)\cdot (l+r))}; \]
\[ [t, i] = \min([t_1, t_2]); \]

%sprintf('either')

else if (N(1,1)==0)

\[ t = \frac{-\log(rand)}{(N(2,1)\cdot (l+r))}; \]
\[ i = 2; \]

%sprintf('type 2')

else if (N(2,1)==0)

\[ t = \frac{-\log(rand)}{(N(1,1)\cdot (l+1))}; \]
\[ i = 1; \]

%sprintf('type 1')

end
end

\% Determine the nature of the latest branching event

d = P(i,1)/(1+P(i)); \% d = death probability

k = 0; \% i.e. default #new infections = 0
if rand < d
    k = -1; % lose an infected individual
else if rand < P(i,2) % i.e. if infection successful
    k = 1;
end
end

% If there is a new infection, check whether it is a mutant
S = [2; 1]; % S for "switch": 1 -> 2, 2 -> 1
if k==1
    if rand < m
        i = S(i);
    end
end

%[k i t]

N = [N(:,1) N]; % Make room for the latest data
    % in the simulated history
N(i,1) = N(i,1)+k; % Update the simulated history
T = [t+T(1) T]; % Update the clock
R = [N(2,1)/N(1,1), R]; % Update the simulated ratios

% Calculate the predicted outcomes for the present time
M = expm([((1-m)*b*l-1 m*b*l; m*g*l (1-m)*g*l-r]*T(1));
O = [O(:,1) O];
O(1,1) = M(1,1); % Predicted #virus 1 from initial type-1 host
O(2,1) = M(1,2); % Predicted #virus 2 from initial type-1 host
%O(3,i) = M(2,1);
%O(4,i) = M(2,2);
%Rp = (O(2,:)+O(4,:))/(O(1,:)+O(3,:));
Rp = O(2,:)./O(1,:); % Predicted ratios of #virus 2 to #virus 1

subplot(2,2,1); plot(T, N)
legend('pathogen 1','pathogen 2','Location','NorthWest')
title(['(Simulated) \alpha = ',num2str(r), '; \beta = ',num2str(b),
    '; \gamma = ',num2str(g), '; \mu = ',num2str(m), '; \lambda = ',num2str(l)])
A = [N; T];

A.2 means

function f = means(r,b,g,m,l,max_time)

% Plot the entries of a specific matrix exponential as a
% parameter changes.
% Input values:
% r: ratio of removal rates (due to death or recovery) bw viruses 1 & 2
% (r>1 => virus 2 is associated with a higher removal rate.)
% b: probability that an encounter with virus 1 -> infection
% g: probability that an encounter with virus 2 -> infection
% m: probability of mutation in a new infection
% l: exposure rate (due to population density, for instance)
% t: maximum time of interest

time = 0:0.1:max_time;
O = [];

for i=1:(10*max_time+1)
    M = expm([(1-m)*b*l-1 m*b*l; m*g*l (1-m)*g*l-r]*time(i));
    O(1,i)=M(1,1);
    O(2,i)=M(1,2);
    O(3,i)=M(2,1);
    O(4,i)=M(2,2);
end

R = O(2,:)./O(1,:);

subplot(2,1,1); plot(time, [O(1,:); O(2,:)]);
legend('Pathogen 1','Pathogen 2')
title(['\alpha = ',num2str(r), '; \beta = ',num2str(b), '; \gamma = ',num2str(g), '; \mu = ',num2str(m), '; \lambda = ',num2str(l)])
xlabel('Time')
ylabel('#infected')

subplot(2,1,2); f=plot(time, R, '-r');
axis([0 max_time max(0,min(R)-1) max(R)+1])
xlabel('Time')
ylabel('Ratio infected by virus 2 to virus 1')

A.3 dsimbranch

This is similar to simbranch, but permits arbitrary initial conditions; this is the code used for subsection 3.2.6.

function A = dsimbranch(r,b,g,m,l, v1, v2)
% A = simbranch(r, b, g, m, l)
%
% Simulates a 2-type cts-time branching process interpreted as a
% competition between viruses in a host population, when there is an initial condition.
%
% Assumes that the initial population has one host infected with virus 1.
%
% Input values:
% r: ratio of removal rates (due to death or recovery) between viruses 1 & 2
% (r>1 => virus 2 is associated with a higher removal rate.)
% b: probability that an encounter with virus 1 -> infection
% g: probability that an encounter with virus 2 -> infection
% m: probability of mutation in a new infection
% l: exposure rate (due to population density, for instance)
% v1: initial number of pathogens type-1
% v2: initial number of pathogens type-2
%
% Outputs a matrix with rows one and two containing the numbers of individuals
% infected by pathogens type 1 and 2 respectively, while the third row
% indexes the split times (from latest in the first column through to
% the initial time.)

N = [v1; v2]; % N will track the number of each type of pathogen in the population.
T = [0]; % T will track the split times
P = [1, b; r, g];
% P stores the pathogen parameters.
% The column entries correspond to removal rate and infection probability,
% the row entries index the pathogens.
R = N(2,1)/N(1,1);
M = [v1 v2; v1 v2]*expm([[(1-m)*b*l-1 m*b*l; m*g*l (1-m)*g*l-r]*T(1));
O = [M(1,1); M(1,2)];

% The following lines of code scale the plot size for best appearance.
% "s" is a parameter more or less corresponding to the percentage of the
% screen you want filled by the plot.
s = .8;
fullscreen = get(0,'ScreenSize');
figure('Position',[(1-s)*fullscreen(3)/2 (1-s)*fullscreen(4)/2 s*fullscreen(3) s*fullscreen(4)],'Name','Simulation vs. Prediction based on 1 initial host with pathogen type 1')
while(any(N(:,1)))
    % The following if/else statements induce the branchings. That is, they
    % generate the split-times for the process.
    if(all(N(:,1)))
        t_1 = -log(rand)/(N(1,1)*(l+1));
        t_2 = -log(rand)/(N(2,1)*(l+r));
        [t, i] = min([t_1, t_2]);
        %sprintf('either')
    else if (N(1,1)==0)
        t = -log(rand)/(N(2,1)*(l+r));
        i = 2;
        %sprintf('type 2')
    else if (N(2,1)==0)
        t = -log(rand)/(N(1,1)*(l+1));
        i = 1;
        %sprintf('type 1')
    end
end
end

% Determine the nature of the latest branching event,
% that is, check whether it is a branch being removed (by death or
% recovery), or whether the split-time corresponds to an infectious
% exposure.
d = P(i,1)/(l+P(i)); % d = death probability
k = 0; % i.e. default #new infections = 0
if rand < d
    k = -1; % lose an infected individual
else if rand < P(i,2)
    % i.e. if infection successful
    k = 1;
end
% If there is a new infection, check whether it is a mutant
S = [2; 1]; % S for "switch": 1 -> 2, 2 -> 1
if k==1
    if rand < m
        i = S(i);
    end
end
%[k i t]
N = [N(:,1) N]; % Make room for the latest data in the simulated history
N(i,1) = N(i,1)+k; % Update the simulated history
T = [t+T(1) T]; % Update the clock
R = [N(2,1)/N(1,1), R]; % Update the simulated ratios

% Calculate the predicted outcomes for the present time
M = [v1 v2; v1 v2]*expm([(1-m)*b*l-1 m*b*l; m*g*l (1-m)*g*l-r]*T(1));
O = [M(:,2)' O];
Rp = O(2,:)./O(1,:); % Predicted ratios of #virus 2 to #virus 1

subplot(2,2,1); plot(T, N)
legend('pathogen 1','pathogen 2','Location','NorthWest')
title(['(Simulated) \alpha = ',num2str(r), '; \beta = ',num2str(b), '; \gamma = ',num2str(g), '; \mu = ',num2str(m), '; \lambda = ',num2str(l)])
axis([0 T(1) 0 max(max(max(N)),max(max(O)))+1])
xlabel('Time')
ylabel('#infected')

subplot(2,2,2); plot(T, O);
legend('pathogen 1','pathogen 2','Location','NorthWest')
axis([0 T(1) 0 max(max(max(N)),max(max(O)))+1])
title(['(Predicted) \alpha = ',num2str(r), '; \beta = ',num2str(b), '; \gamma = ',num2str(g), '; \mu = ',num2str(m), '; \lambda = ',num2str(l)])
xlabel('Time')
ylabel('#infected')

subplot(2,2,3); plot(T, Rp, '-r');
title('Simulated ratio of #virus 2 to #virus 1')
axis([0 T(1) 0 max(max(R),max(Rp))+1])
xlabel('t-t_{C}')
ylabel('Ratio infected by virus 2 to virus 1')

subplot(2,2,4); plot(T, Rp, '-r');
title('Predicted ratio of #virus 2 to #virus 1')
axis([0 T(1) 0 max(max(R),max(Rp))+1])
xlabel('t-t_{C}')
ylabel('Ratio infected by virus 2 to virus 1')

drawnow
end

A = [N; T];

A.4 dmeans

function f = dmeans(r,b,g,m,l,v1,v2,max_time)

% Plot the entries of a specific matrix exponential as a
% parameter changes.
% Input values:
% r: ratio of removal rates (due to death or recovery) bw viruses 1 & 2
% (r>1 => virus 2 is associated with a higher removal rate.)
% b: probability that an encounter with virus 1 -> infection
% g: probability that an encounter with virus 2 -> infection
% m: probability of mutation in a new infection
% l: exposure rate (due to population density, for instance)
% t: maximum time of interest
% v1: initial number of type-1 infecteds
% v2: initial number of type-2 infecteds

time = 0:0.1:max_time;
O = [];

for i=1:(10*max_time+1)
\[ M = \begin{bmatrix} v1 & v2 \\ vl & v2 \end{bmatrix} \expm((1-m)\cdot b_l - 1 \cdot m \cdot b \cdot l - 1 \cdot m \cdot g \cdot l - (1-m) \cdot g \cdot l - r) \cdot \text{time}(i); \]
\[ O(1,i) = M(1,1); \]
\[ O(2,i) = M(1,2); \]
\[ O(3,i) = M(2,1); \]
\[ O(4,i) = M(2,2); \]
end

\[ R = O(2,:)/O(1,:); \]

subplot(2,1,1); plot(time, [O(1,:); O(2,:)]);
legend('Pathogen 1', 'Pathogen 2')
title(['\alpha = ', num2str(r), '; \beta = ', num2str(b), '; \gamma = ', num2str(g), ''])
xlabel('t-t_{C}')
ylabel('#infected')

subplot(2,1,2); f=plot(time, R, '-r');
axis([0 max_time max(0,min(R)-1) max(R)+1])
xlabel('t-t_{C}')
ylabel('Ratio infected by virus 2 to virus 1')
Bibliography


