



UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2019:5

A stochastic differential equation derived from evolutionary game theory

Brian Treacy

Examensarbete i matematik, 15 hp
Handledare: Ingemar Kaj
Examinator: Denis Gaidashev
Februari 2019

A large, faint watermark of the Uppsala University seal is visible in the bottom right corner of the page. The seal features a sun with rays, a cross, and the Latin motto 'ALERE FLAMMAM VERITATIS' around the perimeter.

Department of Mathematics
Uppsala University

Abstract

Evolutionary game theory models population dynamics of species when a concept of fitness is introduced. Initially ordinary differential equation models were the primary tool but these fail to capture inherent randomness in sufficiently small finite populations. Therefore stochastic models were created to capture the stochasticity in these populations. This thesis connects these two modeling paradigms. It does this by defining a stochastic process model for fitness-dependent population dynamics and deriving its associated infinitesimal generator. This is followed by linking the finite population stochastic process model to the infinite population ordinary differential equation model by taking the limit of the total population which results in a stochastic differential equation. The behaviour of this stochastic differential equation is analyzed, showing how certain parameters affect how closely this stochastic differential equation mimics the behaviour of the ordinary differential equation. The thesis concludes by providing graphical evidence that the one-third rule holds with this stochastic differential equation.

Contents

Introduction	5
1 Moran Process	6
1.1 Payoff	6
1.2 Fitness	7
1.3 Transition Probabilities	7
1.4 Fixation Probabilities and Times	8
2 Infinitesimal Generator	9
2.1 Infinitesimal Generator for alternative fitness function	10
3 Stochastic Differential Equations	11
3.1 Heuristics	11
3.2 Existence and Uniqueness	11
3.3 Infinitesimal Generator	12
4 Analysis	13
4.1 Analysis of deterministic component	13
4.2 Analysis of the stochastic component	14
4.2.1 Scaling X_t	14
4.3 Dependency of Q on $p\gamma$	16
4.4 The one-third rule	17
5 Conclusion	19

Introduction

Evolutionary game theory seeks to model the population dynamics of species when a concept of fitness is introduced, specifically that higher fitness corresponds to faster reproduction. A key insight from evolutionary game theory is that an organism's fitness can be heavily dependent on how it interacts with other organisms nearby, be they from its own species or otherwise. Therefore a species' fitness can be a fluid concept that continually changes with the relative proportions of the different species in the population. This frequency-dependent concept of fitness was first mathematically modeled by John Maynard Smith which culminated in his seminal work *Evolution and the Theory of Games* [5]. The field's initial foray consisted of deterministic differential equation models, however these models can be unrealistic in certain situations, as often in the field of biology one models finite populations, whose dynamics have an inherent stochastic element due to the finiteness. Stochastic versions of the original deterministic differential equations were proposed as a possible solution. Particular stochastic process approaches, first conceived in the field of population genetics, were reimaged to help capture these altered dynamics. The development of these different models can be further explored in Martin Nowak's book [6] and in a general survey of the field by Christoph Hauert and Arne Traulsen [4]. This thesis looks at the connection between these models. Specifically the behaviour of the the finite population model as we increase the population limit to infinity and how this corresponds to the deterministic differential equation model of the infinite population. The paper follows the following format:

- Chapter 1: Describing the stochastic process in question: the Moran process.
- Chapter 2: Developing the infinitesimal generator and scaling it to an infinite population.
- Chapter 3: Finding the associated stochastic differential equation with our infinitesimal generator.
- Chapter 4: Analyzing the behaviour of this SDE, comparing it with the infinite population model and verifying the one-third rule.

1 Moran Process

The Moran process, first defined by P.A.P Moran in 1958 [1], is a birth-death process that models the population dynamics of agents from two species, A and B . At each transition in this process, an agent in the population is picked randomly to die and another to clone itself. This way, the total population of agents in the process remains the same. As the rate of death of these agents equals the rate they reproduce at, the lifetime of each agent is exponentially distributed with mean 1. This process, which is a continuous-time Markov Chain $\{X_t\}$, is regular on $[0, 1]$, meaning that every point can be hit from every other point: for every $x \in (0, 1)$ and $y \in (0, 1)$ and $T_y = \inf\{t : X_t = y\}$:

$$\mathbb{P}(T_y < \infty | X_0 = x) = 1 \quad (1)$$

Each birth-death event involves an agent being selected to die randomly and simultaneously an agent being selected to clone itself. The agent picked to clone is also picked randomly but proportionate the fitness of the agent's species. Each species' fitness will be a function of the frequency of each species in the population and a payoff structure that models how all the agents in the population interact with each other.

In order to calculate the transition probabilities for the Moran process, we need to define the payoff structure. This chapter will layout how the transition probabilities of the process are calculated.

1.1 Payoff

We begin specifying the payoff structure for both species; A and B . These payoffs are derived from a two player normal-form game with the following payoff matrix:

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array} \quad (2)$$

Where a, b, c and d are real numbers. The above game is normal-form as it can be represented in matrix form and can be interpreted as the following:

- an agent of species A gets the fixed amount a from interacting with an agent of species A .
- an agent of species A gets the fixed amount b from interacting with an agent of species B .
- an agent of species B gets the fixed amount c from interacting with an agent of species A .
- an agent of species B gets the fixed amount d from interacting with an agent of species B .

Hence, assuming that the agents of each species don't make any conscious decisions with whom they interact with and thus the probability of any two agents from either of the two species interacting with each other is solely a function of the frequency of each species in the population, we define the payoff for each species as:

$$\pi_A = a \frac{i_t - 1}{N - 1} + b \frac{N - i_t}{N - 1} \quad (3)$$

$$\pi_B = c \frac{i_t}{N - 1} + d \frac{N - i_t - 1}{N - 1} \quad (4)$$

In the above definitions for the payoffs, i_t is the number of species A and $N - i_t$ is the number of species B for a fixed total population N . Thus i_t is a nonnegative integer. We also note that self-interactions are excluded (an agent can't interact with itself). However as this paper focuses on the behaviour of this Moran process as N tends to infinity, we see that the minus one terms that account for the exclusion of self-interactions become negligible and hence we remove them for notational simplicity:

$$\pi_A = a \frac{i_t}{N} + b \frac{N - i_t}{N} \quad (5)$$

$$\pi_B = c \frac{i_t}{N} + d \frac{N - i_t}{N} \quad (6)$$

Finally defining $x_t := \frac{i_t}{N}$ to be the proportion of species A at time t , we have our payoffs defined as a function of the proportion of each species:

$$\pi_A = ax_t + b(1 - x_t) \quad (7)$$

$$\pi_B = cx_t + d(1 - x_t) \quad (8)$$

1.2 Fitness

Fitness is important as it influences the probability of an agent being selected to clone itself in the birth-death process. Our fitness will be defined as a convex combination of background fitness (set to 1 below) and the payoff from the game. Let $w \geq 0$ be the intensity of selection. If $w = 0$, then the process becomes an undirected random walk while as $w \rightarrow 1$, then the fitness tends to the payoff from the game.

$$f_A = 1 - w(1 - \pi_A) \quad (9)$$

$$f_B = 1 - w(1 - \pi_B) \quad (10)$$

The above definitions of fitness suffer an issue when the payoffs π_A and π_B are allowed to become negative. As fitness must be positive, due to its role in the transition probability below, negative payoffs mean that there is a maximum intensity of selection. However an alternative fitness function can be used that allows for the intensity of selection w to be any positive constant [7] and results in the same stochastic differential equation being derived as will be shown later. The alternative definitions are:

$$f_A = e^{w\pi_A} \quad (11)$$

$$f_B = e^{w\pi_B} \quad (12)$$

While we will derive the stochastic differential equation for both, assume that we are using the first definitions unless stated otherwise.

1.3 Transition Probabilities

We define the transition probabilities for our Moran process as:

$$x \rightarrow x + \frac{1}{N} \text{ with probability } T_x^+ = \frac{x_t f_A}{x_t f_A + (1 - x_t) f_B} (1 - x_t) \quad (13)$$

$$x \rightarrow x - \frac{1}{N} \text{ with probability } T_x^- = \frac{(1 - x_t) f_B}{x_t f_A + (1 - x_t) f_B} x_t \quad (14)$$

$$x \rightarrow x \text{ with probability } 1 - (T_x^+ + T_x^-) \quad (15)$$

where x is the proportion of agents of population A . To better understand these probabilities, we look at T_x^+ . For $x \rightarrow x + \frac{1}{N}$, we need a member of species A to be selected to duplicate and a member of species B to be selected to die:

$$P(\text{Select } A \text{ to reproduce proportional to fitness}) = \frac{x_t f_A}{x_t f_A + (1 - x_t) f_B} \quad (16)$$

$$P(\text{Select } B \text{ to die}) = 1 - x_t \quad (17)$$

These probabilities are related to each other through the ratio of each species fitness values:

$$T_x^+ = \frac{f_A}{f_B} T_x^- \quad (18)$$

1.4 Fixation Probabilities and Times

Expressions for fixation probabilities and fixation times can be derived. We will state their results here, derivations can be located in Traulsen's and Hauert's review [4]. We define ϕ_j as the probability that j agents of species A takes over and drives species B to extinction. The absorbing states are

$$\phi_0 = 0 \quad \phi_N = 1 \quad (19)$$

Using the absorbing states and iteration techniques we get:

$$\phi_j = \frac{1 + \sum_{k=1}^{j-1} \prod_{i=1}^j \tau_i}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^j \tau_i} \quad (20)$$

where τ_j is the ratio between the transition probabilities $\frac{T_j^-}{T_j^+}$. These transition probabilities are similar to the ones derived before, T_x^- and T_x^+ but differ slightly in that the transition probabilities we derived earlier dealt with the proportions of species A in the population where as T_j^+ and T_j^- are the transition probabilities for when there is j agents of species A in the population. The expressions for T_j^+ and T_j^- can be seen by substituting $\frac{j}{N}$ in for x_t in the expressions derived above for T_x^+ and T_x^- :

$$T_j^+ = \frac{j_t f_A}{j_t f_A + (N - j_t) f_b} \frac{N - j_t}{N} \quad (21)$$

$$T_j^- = \frac{j_t f_A}{j_t f_A + (N - j_t) f_b} \frac{j_t}{N} \quad (22)$$

$$(23)$$

Expressions for the unconditional and conditional fixing times exist as well. Let t_j be the average time until either species A or B fixates, given there are currently j agents of species A and t_j^A be the average time it takes for species A to fixate. These are the unconditional and conditional fixing times respectively and have the following expressions:

$$t_j = -t_1 \sum_{k=j}^{N-1} \prod_{m=1}^k \tau_m + \sum_{k=j}^{N-1} \sum_{l=1}^k \frac{1}{T_l^+} \prod_{m=l+1}^k \tau_m \quad (24)$$

$$t_j^A = -t_1^A \frac{\phi_1}{\phi_j} \sum_{k=j}^{N-1} \prod_{m=1}^k \tau_m + \sum_{k=j}^{N-1} \sum_{l=1}^k \frac{\phi_l}{\phi_j} \frac{1}{T_l^+} \prod_{m=l+1}^k \tau_m \quad (25)$$

2 Infinitesimal Generator

This chapter will focus on deriving the appropriate infinitesimal generator for our defined Moran process. We first begin with a pair of definitions:

A **time homogeneous** continuous-time markov chain is one whose transition probabilities do not explicitly depend on time.

The **infinitesimal generator** of a stochastic process $\{X_t\}_{t \geq 0}$, $X_t \in \{0, \frac{1}{N}, \dots, 1\}$ where X_t is a time homogeneous continuous-time markov chain and $f : \mathbb{R} \rightarrow \mathbb{R}$ where f is twice differentiable is:

$$Lf(x) := \lim_{\delta t \downarrow 0} \frac{E[f(X_{\delta t}) - f(x) | X_0 = x]}{\delta t} \quad (26)$$

As our transition probabilities T_x^+ and T_x^- don't depend explicitly on time, our continuous-time Markov chain X_t is time homogeneous and thus we can use the above definition to derive the infinitesimal generator associated with our process. As $E[f(x) | X_0 = x] = f(x)$, the crux of the issue lies in determining $E[f(X_{\delta t})]$. Defining \mathbb{P}_T to be the probability one transition occurs in the time interval $(0, \delta t]$ and \mathbb{P}_{T_i} to be the probability that i transitions occur in time interval $(0, \delta t]$, we can begin to calculate the expected value:

$$E[f(X_{\delta t}) | X_0 = x] = \mathbb{P}_T [T_x^+ f(x + \frac{1}{N}) + T_x^- f(x - \frac{1}{N}) + (1 - T_x^+ - T_x^-) f(x)] \quad (27)$$

$$+ (1 - \mathbb{P}_T) f(x) \quad (28)$$

$$+ \sum_{i=2}^{\infty} \mathbb{P}_{T_i} g_i(x) \quad (29)$$

where $g_i(x)$ are the permutations of the state space that can occur given i transitions occurred multiplied by the appropriate transition probabilities. Given that the lifetime of each agent is exponentially distributed with mean 1, each birth-death event involves a pair of agents and there are $\binom{N}{2}$ pairs imply that the transitions of this Moran process take place at the points of a Poisson process with rate $\binom{N}{2}$. We can then calculate the probability of k transitions in the the time interval $(0, \delta t]$:

$$\mathbb{P}(|X_{\delta t} - X_0| = k) = \frac{e^{-\delta t \binom{N}{2}} (\delta t \binom{N}{2})^k}{k!} \quad (30)$$

Further, $e^{-\delta t \binom{N}{2}} = 1 - \binom{N}{2} \delta t + O((\delta t)^2)$ by Taylor's theorem. Substituting this back in, we see that

$$\mathbb{P}_T = \mathbb{P}(|X_{\delta t} - X_0| = 1) = \binom{N}{2} \delta t + O((\delta t)^2) \quad (31)$$

and that $\mathbb{P}_{T_k} = O((\delta t)^k)$ for $k \geq 2$. Using Taylors Theorem on $f(x + \frac{1}{N})$ and $f(x - \frac{1}{N})$, we have an expression for our infinitesimal generator:

$$Lf(x) = \binom{N}{2} \left(\frac{1}{N} f'(x) (T_x^+ - T_x^-) + \frac{1}{2N^2} f''(x) (T_x^+ + T_x^-) + O\left(\frac{1}{N^3}\right) \right) \quad (32)$$

Substituting in our expressions for our transition probabilities T_x^+ and T_x^- , we have

$$Lf(x) = \left(\frac{1}{2}\right) \frac{x(1-x)}{xf_A + (1-x)f_B} \left[\left(1 - \frac{1}{N}\right)N(f_A - f_B)f'(x)\right] \quad (33)$$

$$+ \left(\frac{1}{4} - \frac{1}{4N}\right)(f_A + f_B)f''(x) \quad (34)$$

$$+ N(N-1)O\left(\frac{1}{N^3}\right)] \quad (35)$$

Yet this infinitesimal generator faces an issue. As $N \rightarrow \infty$, $Lf(x) \rightarrow \infty$. In particular $N(f_A - f_B)$ blows up. On closer inspection, $f_A - f_B = w(\pi_A - \pi_B)$ and thus in order for the limit to behave nicely, we must define a new scale of intensity w_N such that $Nw_N = \gamma$, for some $\gamma > 0$. γ must be positive as the intensity of selection parameter w_N is defined positive. Thus we have:

$$w_N = \frac{\gamma}{N} \quad (36)$$

$$f_A - f_B = w_N((a-c)x + (b-d)(1-x)) \quad (37)$$

$$f_A + f_B = 2 - 2w_N + w_N((a+c)x + (b+d)(1-x)) \quad (38)$$

Substituting the above into the expression for our infinitesimal generator $Lf(x)$ and letting $N \rightarrow \infty$:

$$Lf(x) = \frac{\gamma}{2}x(1-x)[(a-c)x + (b-d)(1-x)]f'(x) \quad (39)$$

$$+ \frac{1}{4}x(1-x)f''(x) \quad (40)$$

2.1 Infinitesimal Generator for alternative fitness function

We mentioned previously that an alternative fitness definition could be used. This alternative definition is $f_A = e^{w_N\pi_A}$ and $f_B = e^{w_N\pi_B}$. With this definition the only difference in defining the infinitesimal generator lies in scaling the selection parameter with the term $N(f_A - f_B)$ that goes to infinity as $N \rightarrow \infty$. However we can Taylor expand $(f_A - f_B)$ to address this:

$$(f_A - f_B) = e^{w_N\pi_A} - e^{w_N\pi_B} \quad (41)$$

$$= w_N(\pi_A - \pi_B) + w_N^2 \frac{\pi_A^2 - \pi_B^2}{2} + O(w_N^3) \quad (42)$$

$$= w_N(\pi_A - \pi_B) + O(w_N^2) \quad (43)$$

Given $w_N = \frac{\gamma}{N} \rightarrow 0$ as $N \rightarrow \infty$, this leaves us with the exact same expression for $N(f_A - f_B)$ in the limit of N as the original definition and hence leaves us with the same infinitesimal generator $Lf(x)$.

3 Stochastic Differential Equations

To find the stochastic differential equation associated with our Moran process, we first employ a heuristic method and then justify the results.

3.1 Heuristics

We begin by assuming the infinitesimal generator derived in the previous chapter is described by a stochastic differential equation:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (44)$$

where W_t is a one-dimensional Brownian Motion. This assumption will be justified later. Let $f : \mathbb{R} \rightarrow \mathbb{R}$, such that f is twice differentiable. We then apply Itô's Lemma:

$$df(X_t) = [\mu(X_t)f'(X_t) + \frac{1}{2}\sigma^2(X_t)f''(X_t)]dt + \sigma(X_t)f'(X_t)dW_t \quad (45)$$

Comparing this with our infinitesimal generator, we can select our values for μ and σ by comparing the coefficients of the $f'(x)$ and $f''(x)$ terms:

$$Lf(x) = \frac{\gamma}{2}x(1-x)[(a-c)x + (b-d)(1-x)]f'(x) + \frac{1}{2}x(1-x)f''(x) \quad (46)$$

$$\mu(X_t) = \frac{\gamma}{2}X_t(1-X_t)[(a-c)X_t + (b-d)(1-X_t)] \quad (47)$$

$$\sigma(X_t) = \frac{1}{2}\sqrt{X_t(1-X_t)} \quad (48)$$

Having found this SDE, we must now work backwards. We will show that the solution to this SDE exists and is unique. Then we will show that it has an associated infinitesimal generator, which will be the one derived in the previous chapter.

3.2 Existence and Uniqueness

The proof for existence requires considerable amounts of measure theory and can be sourced in Revuz and Yor [3], Theorem 9.1.7. To prove uniqueness, we rely on theorem 24.2 from Bass [9] that states if $\mu(X_t)$ is Lipschitz and bounded and that there exists $\rho : [0, \infty) \rightarrow [0, \infty)$ such that $\rho(0) = 0$,

$$\int_0^\varepsilon \rho^{-2}(u)du = 0 \quad (49)$$

for all $\varepsilon > 0$, and σ is bounded and satisfies

$$|\sigma(x) - \sigma(y)| \leq \rho(|x - y|) \quad (50)$$

for all x and y , then the solution to the SDE in question is pathwise unique. We will now show that the above conditions are satisfied and a suitable ρ function exist.

$\mu(X_t)$ is bounded: $\mu'(x) = \frac{\gamma}{2}(-2px^2 + (p-q)x + 1)$ so therefore there exists x_1 and x_2 such that $\mu'(x_1) = \mu'(x_2) = 0$. This means one of $\mu(0), \mu(1), \mu(x_1)$ and $\mu(x_2)$ is an upper bound for $\mu(x)$ on the domain $(0, 1)$. As $\mu(x) = \frac{\gamma}{2}x(1-x)(px + q)$ has no singularities, we can be sure that $\mu(x)$ is bounded.

$\mu(X_t)$ is Lipschitz:

$$\mu(x) - \mu(y) = \frac{\gamma}{2}(x^2(1-x)(a-c) + x(1-x)^2(b-d) - y^2(1-y)(a-c) - y(1-y)^2(b-d)) \quad (51)$$

$$= \frac{\gamma}{2}(x-y)((a-c)(x+y-x^2-xy-y^2) + (b-d)(1-2(x+y)+x^2+xy+y^2)) \quad (52)$$

$$= (x-y)g(x,y) \quad (53)$$

As both x and y are elements of the set $[0, 1]$ and $g(x, y)$ is a polynomial of order 2, there exists a constant $c \in [0, 1]$ such that $|g(x, y)| \leq c$. Therefore

$$\frac{|u(x) - u(y)|}{|x - y|} = \frac{|x - y| |g(x, y)|}{|x - y|} \quad (54)$$

$$= |g(x, y)| \leq c \quad (55)$$

Therefore $\mu(x)$ is Lipschitz.

$\sigma(X_t)$ is bounded: As $\sigma(x) = \frac{1}{2}\sqrt{x(1-x)}$, $\sigma(x)$ is bounded by $\sigma(\frac{1}{2})$.

$$\rho(x) = \sqrt{x}$$

- $\rho(0) = 0$
- $\int_0^\varepsilon \rho^{-2}(u) du = \ln(\varepsilon) - \ln(0) = \infty$.

Finally for our values of $x, y \in [0, 1]$, the required inequality was tested to be true in Mathematica. Hence, the solution to the SDE is unique.

3.3 Infinitesimal Generator

Finally, with existence and uniqueness shown, we must justify our heuristic logic used in the beginning of this chapter; connecting our stochastic differential equation to the infinitesimal generator. We rely on theorem 39.3 from Bass [9]: For the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (56)$$

where σ and μ are both Borel-measurable and bounded, then for a function $f \in C^2$, the solution to that stochastic differential equation X_t satisfies

$$f(X_t) = f(X_0) + \int_0^t \frac{\partial f}{\partial x}(X_s)\sigma(X_s)dW_s + \int_0^t Lf(X_s)ds \quad (57)$$

where $Lf(x)$ is defined as

$$Lf(x) = \frac{1}{2}\sigma(x)\frac{\partial^2 f}{\partial x^2} + \mu(x)\frac{\partial f}{\partial x} \quad (58)$$

Using the values we derived for σ and μ , we get the original infinitesimal generator.

4 Analysis

In this chapter we compare the behaviour of the derived stochastic differential equation with the behaviour of the corresponding deterministic differential equation.

4.1 Analysis of deterministic component

For the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \quad (59)$$

where μ and σ are defined as

$$\mu(X_t) = \frac{\gamma}{2} X_t(1 - X_t)[(a - c)X_t + (b - d)(1 - X_t)] \quad (60)$$

$$\sigma(X_t) = \frac{1}{2} \sqrt{X_t(1 - X_t)} \quad (61)$$

we see there is close alignment between $\mu(X_t)$ and the deterministic differential equation model below:

$$\dot{x} = x(1 - x)[(a - c)x + (b - d)(1 - x)] \quad (62)$$

This deterministic differential equation has four generic cases:

- **Dominance** This case results in one species driving the other to extinction. *A* drives *B* to extinction, which we refer to as *A* fixating, if $a > c$ and $b > d$. In this scenario there are only two fixed points, $x = 1$ is a stable fixed point and $x = 0$ is unstable. The other possible dominance case is *B* fixating, which happens if $a < c$ and $b < d$. In this case, $x = 1$ is an unstable fixed point and $x = 0$ is stable.
- **Bistability** occurs when $a > c$ and $b < d$. There are three fixed points in this case: $x = 0$ and $x = 1$ are stable where as the middle fixed point $x^* = \frac{-(b-d)}{a-b-c+d}$ is unstable.
- **Coexistence** occurs when $a < c$ and $b > d$, three fixed points again: $x = 0$ and $x = 1$ are unstable where as x^* is stable, where x^* is defined as in the Bistability case.
- **Neutrality** occurs when $a = c$ and $b = d$. Neutrally stable fixed points for all x .

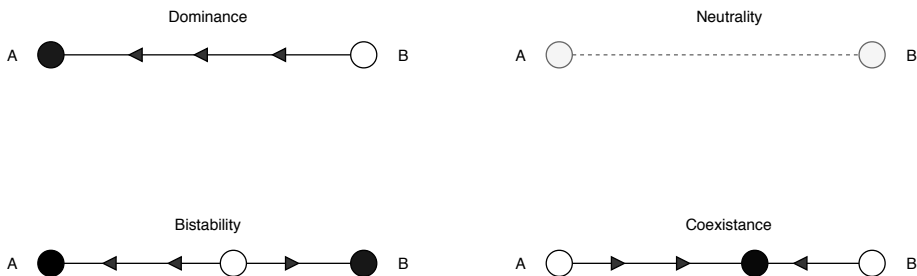


Figure 1 – The four generic cases showing stable (black) and unstable (white) fixed points. The left-most point labeled A indicates the species A has fixated and likewise for B with the rightmost point. The arrows indicate the 'direction of selection'.

4.2 Analysis of the stochastic component

For analysis of the stochastic influence on the evolution of our process, we will focus on deriving $Q(x)$, which will be the probability that the species A drives the species B to extinction. Therefore, defining $T_i = \inf\{t > 0 : X_t = i\}$, our expression for $Q(x)$ is:

$$Q(x) = \mathbb{P}(T_1 < T_0 | X_0 = x) \quad (63)$$

4.2.1 Scaling X_t

If X_t is a standard Brownian motion, then proposition 3.16 of Bass [9] states the distribution of X_t upon exiting an interval $[I_1, I_2]$ is

$$\mathbb{P}(T_{I_1} < T_{I_2} | X_0 = x) = \frac{I_2 - x}{I_2 - I_1} \quad \mathbb{P}(T_{I_2} < T_{I_1} | X_0 = x) = \frac{x - I_1}{I_2 - I_1} \quad (64)$$

If X_t is a regular diffusion such that the property above holds for every interval $[a, b]$, then we say that X_t is on natural scale. If our regular diffusion X_t isn't on natural scale, then it is possible to find a continuous, strictly increasing scale function $S(x)$ such that $S(X_t)$ is on natural scale. For the SDE $dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$ where both σ and μ are real-valued, continuous and bounded above and σ is bounded below by a positive constant, then theorem 41.1 in Bass [9] states that the scale function $S(x)$ is given below, where c_1, c_2 , and x_0 are some constants:

$$S(x) = c_1 + c_2 \int_{x_0}^x e^{-2 \int_{x_0}^y \frac{\mu(z)}{\sigma^2(z)} dz} dy \quad (65)$$

To simplify the calculations, we define

$$\begin{aligned} p &= a - c - b + d \\ q &= b - d \end{aligned}$$

Rewriting $\mu(x)$:

$$\mu(x) = \frac{\gamma}{2} x(1-x)(px+q)$$

Substituting in $\mu(z)$ and $\sigma(z)$ to the scale function $S(x)$:

$$S(x) = c_1 + c_2 e^{2\gamma p(x_0 + \frac{q}{p})^2} \int_{x_0}^x e^{-2\gamma p(y + \frac{q}{p})^2} dy \quad (66)$$

Now that $S(X_t)$ is in natural scale, we then determine the distribution of species A driving species B to extinction and fixating:

$$Q(x) = \frac{S(x) - S(0)}{S(1) - S(0)} \quad (67)$$

$$= \frac{\int_0^x e^{-2\gamma p(y + \frac{q}{p})^2} dy}{\int_0^1 e^{-2\gamma p(y + \frac{q}{p})^2} dy} \quad (68)$$

Remembering that two of the four generic cases in the analysis of the deterministic component had a third fixed point, the other two being $x = 0$ and $x = 1$, we see the expression

for this fixed point emerge in our expression for $Q(x)$:

$$x^* = \frac{-(b-d)}{a-c-b+d} \tag{69}$$

$$= \frac{-q}{p} \tag{70}$$

Substituting this into $Q(x)$, we can manipulate $Q(x)$ into being the ratio between two slivers of a normal distribution with mean x^* and variance $\frac{1}{4\gamma p}$:

$$Q(x) = \frac{\int_0^x \frac{1}{\sqrt{2\pi\frac{1}{4\gamma p}}} e^{-2\gamma p(y-x^*)^2} dy}{\int_0^1 \frac{1}{\sqrt{2\pi\frac{1}{4\gamma p}}} e^{-2\gamma p(y-x^*)^2} dy} \tag{71}$$

$$\tag{72}$$

The behaviour of $Q(x)$ is captured by the shape of the distribution over the interval $[0, 1]$. As γp increases, the variance of the distribution decreases, thereby flattening the graph of $Q(x)$. However the centering of the distribution matters as well. A very steep distribution might still have a rather flat $Q(x)$ graph if the center of the distribution x^* is sufficiently far away from the interval $[0, 1]$. $Q(x)$ has a nice visual explanation as the ratio of two areas of a normal distribution:

$$Q(x) = \frac{\text{Light Blue Area}}{\text{Light Blue Area} + \text{Dark Blue Area}} \tag{73}$$

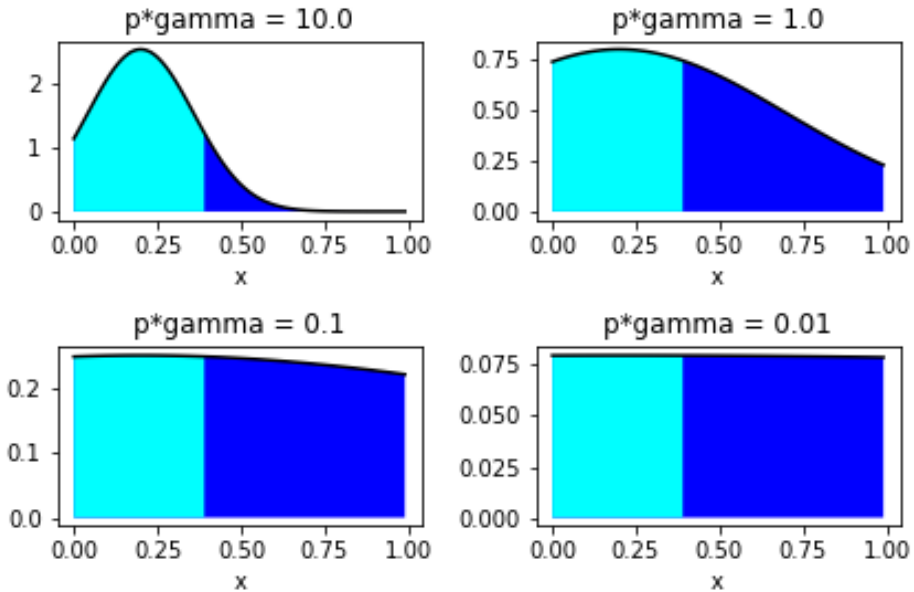


Figure 2 - Altering p can change the shape of the distribution, thus changing the area of the light blue area respective the the area of the dark blue area

4.3 Dependency of Q on $p\gamma$

Recalling γ was a constant constructed such that that our selection pressure constant would scale as $N \rightarrow \infty$, we see it acts with p and q to scale the deterministic part of the SDE:

$$dX_t = \frac{1}{2}X_t(1-X_t)(\gamma pX_t + \gamma q) + \frac{1}{2}\sqrt{X_t(1-X_t)} \quad (74)$$

We can see this influence visually on the function $Q(x)$. Pairs of p and q were generated randomly and then were grouped by the four generic cases (remembering that p and q are functions of the payoff parameters a, b, c and d which control the case). By fixing the value of γ , we can then see the influence of increasing the magnitude of $p\gamma$ on $Q(x)$. We see that as the magnitude of $p\gamma$ increases, the graph for $Q(x)$ comes to reflect the case for each of the four generic cases, as seen in figure 3. However as the magnitude of $p\gamma$ decreases, the graph of $Q(x)$ reflects a random walk, as seen in figure 4.

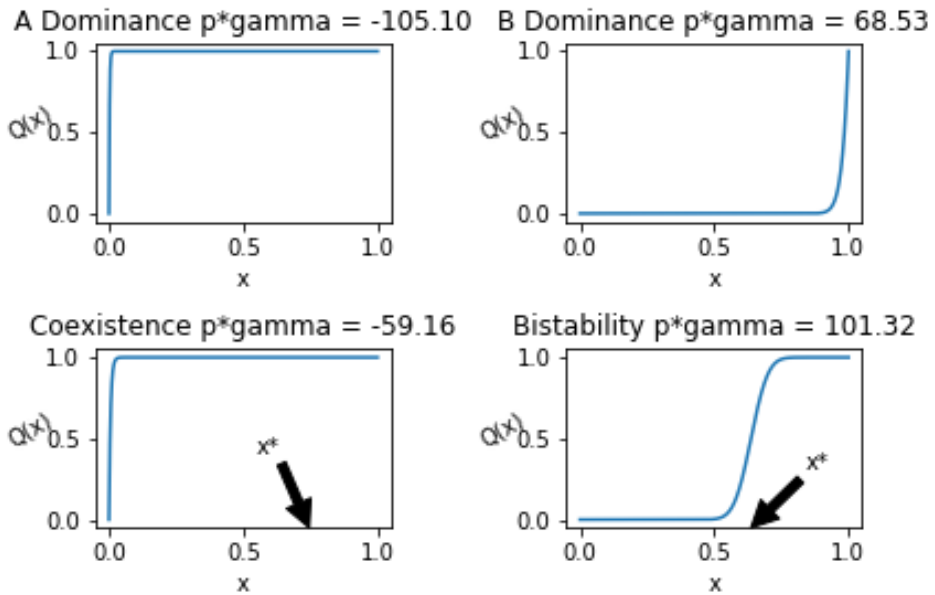


Figure 3 - For high values of $p\gamma$, we see that the graph of $Q(x)$ adheres to the behaviour of the generic case strongly. Note the coexistence case switches between mimicking A dominance and B dominance based on where the fixed point is. The closer the fixed point is to 1, then likelihood that the process will fixate with species A becomes near certain.

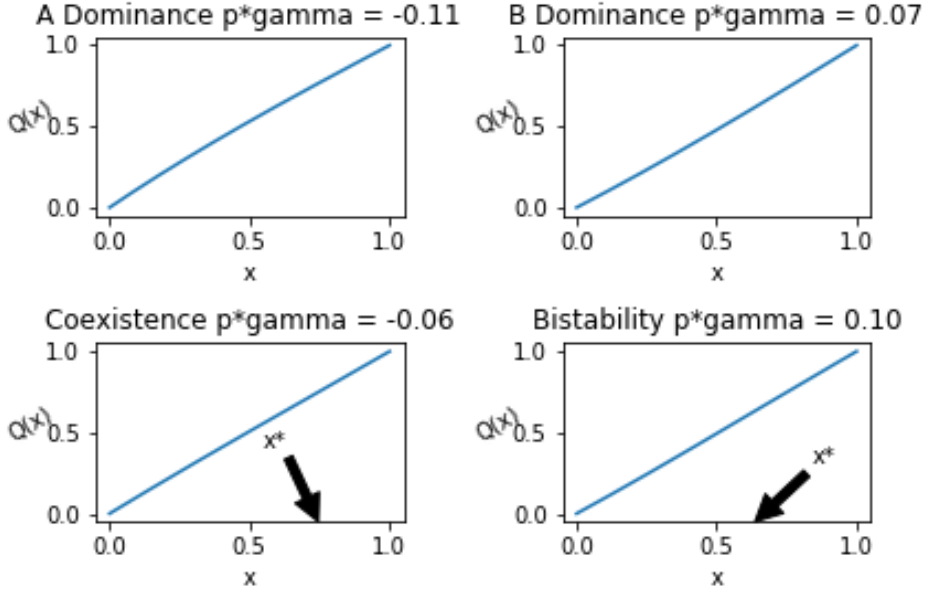


Figure 4 - For low values of $p\gamma$, we see that the graph of $Q(x)$ is more strongly influenced by the stochastic element of the SDE.

4.4 The one-third rule

The one-third rule "establishes the conditions, in the limit of weak selection and large population size under which one Nash Strategy can be invaded by another" [8]. In the context of the system described in this paper, the one-third rule states that the probability of an infinitesimally small number of species A invading and replacing a population of species B in the bistability case is greater than the probability of the same event happening in the neutrality case. $Q(x)$ and x are the probabilities of A fixating in the bistability and neutrality cases respectively. We want to show the following:

$$\lim_{x \rightarrow 0} Q(x) > \lim_{x \rightarrow 0} x \quad (75)$$

However as $Q(0) = 0$, it is enough to show that the derivative of $Q(x)$ is greater than the derivative of x at $x = 0$:

$$Q'(x) - 1|_{x=0} > 0 \quad (76)$$

Remembering that $x^* = -\frac{q}{p}$ which implies that $q = -px^*$

$$Q'(x)|_{x=0} = \frac{1}{\int_0^1 e^{-2\gamma p y^2 + 4\gamma p x^* y} dy} = \frac{1}{A(\gamma, p)} \quad (77)$$

Therefore

$$Q'(x) - 1|_{x=0} > 0 \implies A(z) - 1 < 0 \quad (78)$$

where $z = \gamma p$. Therefore if the one-third rule is true, the above inequality only holds when $x^* < \frac{1}{3}$. Defining $H(z) := A(z) - 1$, we find evidence for the one third rule graphically, as seen in figure 5.

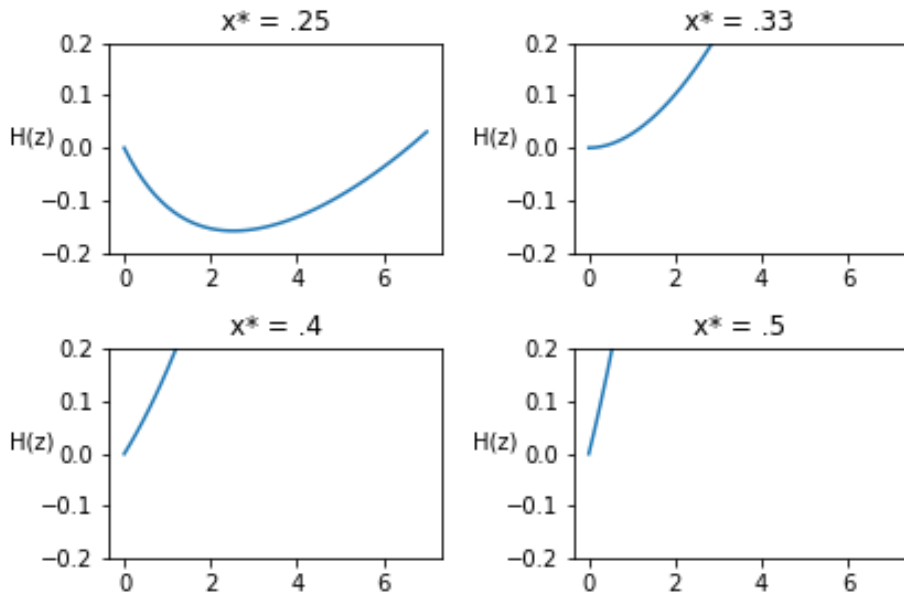


Figure 5 - We see that $H(z)$ only becomes negative when $x^* < \frac{1}{3}$, providing graphical evidence for the one-third rule.

5 Conclusion

When modeling a problem, we can frequently choose from a range of models that will aim to capture different aspects of the problem. Often it is useful to compare these models as similarities between the models assure us that our line of thinking is clear and that our understanding of the problem is strong while differences between the models reveal how the various models capture the distinct nuances of the problem. This thesis looked at two modeling paradigms that were close enough to find reassuring similarities while different enough to pose interesting research questions. The main modeling paradigm followed was a stochastic process approach. Following previous work, we modeled the finite population dynamics of two species whose intra-species and inter-species interactions affected their reproduction. This was done in a stochastic processes framework, specifically using a Moran process, which captured the stochasticity arising from continually shifting demographics. However there are other ways one can view this problem, such as using a deterministic infinite-population ordinary differential equation model. Thus the natural question that arises is how these models relate to each other? To illuminate the similarities and differences between these models, we investigated scaling the total population size in our finite population model.

The stochastic differential equation derived from scaling our stochastic process model showed structural similarities to the deterministic infinite population model, mirroring its behaviour when adjusting certain parameters. Crucially there was also differences due to the stochasticity. One difference was the one-third rule. The one-third rule states that the probability of an infinitesimal amount of one species invading and replacing a population of the other species in the bistability case is greater than the probability of the same event happening in the baseline neutrality case. This rule is of interest as it is impossible in deterministic model and thus is a demonstration of where these modeling paradigms diverge. This thesis is not alone in studying the behaviour of scaled stochastic models. Others have utilized master equations and derived Fokker-Plank equations [10]. This paper's approach of forming an infinitesimal generator was inspired from resources in mathematical genetics [11] and hopefully complements the rest of the research literature well.

References

- [1] P.A.P Moran, *Random processes in genetics*, Mathematical Proceedings of the Cambridge Philosophical Society, Volume 54, Issue 1, January 1958 , pp. 60-71, <https://doi.org/10.1017/S0305004100033193>
- [2] Leslie Lamport, *TeX: a document preparation system*, Addison Wesley, Massachusetts, 2nd edition, 1994.
- [3] Revuz, D. Yor, M. *Continuous Martingales and Brownian Motion*, Grundlehren der mathematischen Wissenschaften, 1999.
- [4] A. Traulsen and C. Hauert, *Stochastic Evolutionary Game Dynamics*, Reviews of Nonlinear Dynamics and Complexity, June 2010. <https://doi.org/10.1002/9783527628001.ch2>
- [5] J.M. Smith, *Evolution and the Theory of Games*, Cambridge University Press, 1982.
- [6] M.A. Nowak, *Evolutionary Dynamics - Exploring the Equations of Life*, The Belknap Press of Harvard University Press, 2006.
- [7] A. Traulsen, N. Shores and M. Nowak, *Analytical Results for Individual and Group Selection of Any Intensity*, Bulletin of Mathematical Biology, April 2008, doi:10.1007/s11538-008-9305-6
- [8] A. Traulsen, J. M. Pacheco, and M. A. Nowak. , *Pairwise comparison and selection temperature in evolutionary game dynamics.*, J. Theor. Biol., 2007, doi:246:522–529,
- [9] R.F. Bass, *Stochastic Processes*, Cambridge Series in Statistical and Probabilistic Mathematics, 2011.
- [10] A. Traulsen, J. C. Claussen, and C. Hauert , *Coevolutionary Dynamics: From Finite to Infinite Populations*, Phys. Rev. Lett. 95, 238701, 2005, <https://doi.org/10.1103/PhysRevLett.95.238701>
- [11] A.M. Etheridge, *Diffusion Process Models in Mathematical* <http://www.stats.ox.ac.uk/etheridg/>