

EXTINCTION, FIXATION, AND INVASION IN AN ECOLOGICAL NICHE

by

MattheW Badali

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

© Copyright 2018 by MattheW Badali

Abstract

Extinction, Fixation, and Invasion in an Ecological Niche

MattheW Badali

Doctor of Philosophy

Graduate Department of Physics

University of Toronto

2018

Contents

1	Ch0-Introduction	1
1.1	Introduction	1
1.2	Motivation and background and such	2
1.2.1	Biodiversity	2
1.3	Neutrality	3
1.3.1	Moran and other simple stochastic models	3
1.4	Niches	4
1.4.1	Concept of a niche, and the debates therein	4
1.4.2	Lotka-Volterra	5
1.5	Stochastics	6
1.5.1	introduction	6
1.5.2	Extinction rates from demographic and environmental stochasticity	7
1.5.3	Approximation techniques	7
1.6	Structure of remaining thesis	7
2	Ch3-AsymmetricLogistic	9
2.1	Invasion Analysis	9
2.2	Discussion	12
2.3	Introduction	12
2.4	Moran Model in More Detail	12
2.5	Moran With Immigration	14
2.6	Some Results	19
	Bibliography	21

Chapter 1

Ch0-Introduction

1.1 Introduction

This thesis is concerned with demographic stochastics. That is, the randomness inherent in systems with a discrete state space. In biology this arises naturally in ecological systems. Stochastics, as applied in the biological context, was first done by Kimura when calculating the dynamics of gene frequencies in a population. Kimura, and most theoretical ecologists since, employed the Fokker-Planck equation, a partial differential equations method which further approximates the system by assuming continuous population sizes. In the context of population ecology, the similar Moran model of two species is the cleanest example of two competing species in an ecosystem, with eventually one going extinct and one fixating after some short characteristic time dependent on the system size. However, this model assumes the two species are identical, and that they compete with each other (interspecies) as strongly as they compete with themselves (intraspecies). Some recent researchers have addressed this by noticing that results similar to that of Moran are found in one limit of the famous generalized Lotka-Volterra equations with stochastic fluctuations. They employ various approximate techniques, usually the Fokker-Planck equation, and explore various metrics of this noisy Lotka-Volterra model, which in other limits has a long average extinction time. None, however, have looked at how the system transitions between its slow and fast limits. Most have also restricted themselves to uncontrolled approximations. This is where I situate my research. In this thesis I will show that competing species can coexist unless their ecological niches entirely overlap, and that this niche overlap anticorrelates with a species' ability to invade an established ecosystem. To accomplish this I shall first perform a thorough investigation of the various approximation techniques commonly used in stochastic biophysical modelling on a one dimensional Lotka-Volterra toy model. Thence I will investigate the two dimensional version, in particular to characterize the transition between its regular slow dynamics and the fast times limit corresponding to the foundational Moran model. I do this with an arbitrarily accurate technique for calculating mean fixation times. Finally, by regarding the opposite process to fixation, I can comment on the stability of the 2D model with regards to invasion attempts. The obvious consequence of my research is a better null hypothesis for the dynamics of small homogeneous communities like the human microbiomes. More generally the results are of significance in estimations of timescale for paleontology and phylogeny.

1.2 Motivation and background and such

1.2.1 Biodiversity

In biology there is a law, or principle, named for Gause [?], which states that “two species cannot coexist if they share a single [ecological] niche.” This is better known as the competitive exclusion principle. That is, in systems with few resources and therefore few niches, one expects that only few species will persist at any given time. But this is not what is observed in nature. Hutchinson outlined the problem with his famous paradox of the plankton [9]; in the top layer of the open ocean there are only a few energy sources and very few minerals or vitamins, yet the number of different phytoplankton living in what seems like the same environment is astounding. The expectation is that in this homogeneous ecosystem with extreme nutrient deficiency the competition should be severe, and only a few species should persist, many fewer than the number observed. A variety of solutions have been proposed but there is as yet no consensus [18].

The problem has persisted for more than half a century, and people continue to research the more general problem of biodiversity and its causes [4, 14, 17?]. Sometimes the research question is complicated, manifesting itself as a difficulty in describing the origin of species abundance distributions. The development of Hubbell’s neutral theory was motivated to explain observed abundance distributions [8]. It contrasts with niche theories of resource apportionment; whereas the former assumes that all species compete with each other, the latter assumes that each species grows based on the apportionment it is allocated and does not touch the resources of other species. Problems in biodiversity can be simpler. One question this text asks is how long a single species is expected to survive, given favourable conditions [?]. Much research has been done on two species competing with each other, as a reduction of the full problem of biodiversity [?]. Whether two species will coexist, and for how long, is of essential importance to the larger problem of biodiversity.

The theory has many applications. Most obvious, and most pressing to society, is the realm of conservation biology. Biodiversity is often used as an indicator of the health of an ecosystem. A clearer understanding of the forces that maintain biodiversity could provide new and easier metrics for evaluating the health of an ecosystem, and hence the efficacy of various conservation efforts. The mechanisms of species maintenance are related to those of speciation, and stability of an ecosystem can refer to both avoiding collapse and avoiding invasion. Invasion of a new mutant or immigrant strain or species into the system is a problem deeply intertwined with that of biodiversity maintenance. This problem too is of obvious interest in the study of ecosystems. Invasion is also relevant in the healthcare world. We are only recently learning, for example, about the composition of the gut microbiome in humans and its relation to health. The balance of different species in ones gut seems to be important for avoiding illness. Imbalance of the microbiome composition, or invasion of a new species, can greatly impact a person’s wellbeing, and a theory of whether an invasion will be successful and how long it might persist would go a long way toward diagnostics and prognostication. The other end of the process, namely the extinction of a species, also has a number of applications. Other than the obvious modern ecological ones, extinction times are useful in paleontology. The fossil record shows a number of species in different epochs, and these data make more sense in the light of a consistent theory of species survival and eventual decline. Similarly, extinction and fixation times are already used in the construction of phylogenetic trees. The more accurate a theory of extinction timescales developed, the more precise we can perform phylogenetic analyses.

1.3 Neutrality

1.3.1 Moran and other simple stochastic models

Start with a simple model of fixation with 2 species, for which we can calculate the time to one species taking over the system. The Moran model [15] is such a model, a classic urn model used in population dynamics in a variety of ways. Its most prominent use is in coalescent theory, describing how the relative proportion of genes in a gene pool might change over time. But really it can describe any system where individuals of different species or strains undergo strong but unselective competition in some closed or finite ecosystem.

To arrive at the Moran model we must make some assumptions. Whether these are justified depends on the situation being regarded. The first assumption is that no individual is better than any other; that is, whether an individual reproduces or dies is independent of its species and the state of the system. This makes the Moran model a neutral theory, and any evolution of the system comes from chance rather than from selection.

Next we assume that the the population size is fixed, owing to the (assumed) strict competition in the system. That is, every time there is a birth the system becomes too crowded and a death follows immediately. Alternately, upon death there is a free space in the system that is filled by a subsequent birth. In the classic Moran model each pair of birth and death event occurs at a discrete time step (cf. the Wright-Fisher model, where each step involves N of these events). This assumption of discrete time can be relaxed without a qualitative change in results.

In the Moran model, each step there is a birth and a death. A species is chosen for either according to its frequency. There is an equal net rate of change, in both increasing and decreasing the frequency. The system fluctuates until either the species dies (extinction) or all others die (fixation). Both of these cases are absorbing states, so once the system reaches either it will never change. Since a species is equally likely to increase or decrease each time step, the model is akin to an unbiased random walk, and therefore the probability of extinction occurring before fixation is known. In this system we can define the first passage time as the time the system takes to reach either fixation *or* extinction. Its calculation is also known.

Under the approximation of continuous population fraction the Moran model effectively becomes that of Kimura [?]. Kimura was inspired by alleles rather than species, but the rationale is similar. In each generation each organism provides many copies of its genome, which are chosen indiscriminately (because each organism has two copies of its genome, a factor of two shows up in Kimura's results when compared those of Moran). Following a few assumptions, Kimura calculates the new mean and variance of the system, hence generating a diffusion equation. Applying the Fokker-Planck approximation to the Moran model obtains the same equation, hence the claim that Kimura's results are similar to those of Moran. Kimura's model similarly can be modified to include many biological effects, like selection. The works of Kimura are well-respected and highly motivated a change in biology to be more quantitative and predictive. Most of Kimura's predictions are numerical by necessity, as no nice analytic forms exist for the solutions. Furthermore, transient behaviour was especially difficult to capture in the models, so only steady states are regarded. Nevertheless, Kimura's ground-breaking work is powerful and wide-ranging. Chapter 4 of this thesis compares some of its outcomes to those from a Kimura paper published decades earlier. His legacy is inescapable.

The seminal work of Hubbell [8] is also similar to that of Moran, but Hubbell is a much more

controversial figure than Kimura. Hubbell, like Moran, was concerned with species, but did not limit himself to Moran’s pedagogical choice of two. Hubbell assumes that each organism from any species competes equally with all others, and therefore as with Moran its probability of reproducing or dying is proportional to its fraction of the population. But Hubbell does not predict fixation probabilities and times. Rather, he calculates the distribution of species abundances that should be present within his neutral model. Following the arguments of Hubbell, one can get an estimate of the expected biodiversity of a community. The abundance distribution he predicts matches well with experimental observations in a variety of biological contexts, from trees to birds to microbiomes. Nevertheless, Hubbell’s neutral theory is contentious. The idea that each species competes with all others to the same degree as intraspecies competition strains credibility. Surely the differences between species matters! Of course there are differences between species; even the staunchest neutralist would agree. But slight perturbations from Hubbell’s theory do not significantly alter its results. What’s more, while everyone concedes that there are differences between species, some argue that these differences do not matter. In some sense, they claim, the species are equivalent and behave neutrally, which is why Hubbell’s theory seems to work so well in such disparate ecologies. The examples presented in Hubbell’s seminal book are compelling, and there may be some truth to these claims. The other side of the debate insists that species differences are the cause of observed abundance distributions. In particular, the environment can be divided into various ecological niches, and it is how these niches are uniquely used by their occupying species that determines the biodiversity of the ecosystem. This is broadly known as niche theory. Niche theory itself has a contentious past mired by confusions. I will do my best to provide a summary here, to contextualize my research.

1.4 Niches

1.4.1 Concept of a niche, and the debates therein

Of course species *aren’t* the same as each other. Some would live happily as the only animals on an island, and others would die out in such a situation. Some can aerobically digest citrate, and others cannot. This is the domain of the competitive exclusion principle. In any given niche, one species will eventually dominate (and usually this is the species optimized to that niche, though this is not necessary for the definition of Gause’ law). This begs the question, what is an ecological niche?

On the theory of niches, Hutchinson [?] says, “Just *because* the theory is analytically true and in a certain sense tautological, we can trust it in the work of trying to find out what has happened” to allow for coexistence of species. In principle, species coexist because they inhabit different niches. The concept of niches is an old one, over a century old, and was popularized by Grinnell [?]. There is therefore over a century of debate as to the meaning of a niche, as there is ambiguity in its use. Following Leibold [11], I refer to the definition of a niche according to the two major camps as the habitat or requirement niche and the functional or impact niche.

Grinnell [?] refers to those environmental considerations that a species can live with as what defines the niche. These include those organisms on different trophic levels than the species, like their predators and prey, but not those on the same trophic level that might compete with them. Hutchinson [?] was in the same camp as Grinnell, and has provided one of the most enduring conceptualizations of a niche, that of an “ n -dimensional hypervolume” in the space of factors that could affect the growth or death of

a species. For each factor there is some range at which the species can reproduce faster than it dies out. This is true both for abiotic factors such as temperature, and biotic factors like the concentration of predators. Sometimes these ranges are bounded by zero, sometimes they are unbounded, and sometimes they depend on the values of the other factors involved. But in the space of all these factors, Hutchinson calls the fundamental niche that the volume in which the species would have a greater birth rate than death rate. He defines the realized niche as the point or subspace in this high dimensional space that the species effectively experiences, given that it is existing and potentially coexisting in an ecosystem. This also lends a natural definition of niche overlap, as the (normalized) overlap of the fundamental niches of two species. A simple model of two species suggests that the functional niche tells us whether the coexistence point of two species is physical [?].

The other usage of the term niche was popularized by Elton [?] and MacArthur & Levins [13], that of a functional or impact niche. Whereas the requirement niche focuses on what factors a species needs to live, the impact niche looks at how the species affects these factors. Their conception of a niche describes how a species influences its environment, or how that species fits in a food web; essentially, what its role is in an ecosystem. This idea is especially attractive to those who study keystone species, but is intuitively understood by anyone who has surveyed a variety of ecosystems. In every ecosystem with flowers there is something that pollinates them; in every ecosystem with cells that grow cellulose cell walls there is something that can digest that cellulose; in every system with prey there are predators. Whether the pollinator is a bird or any number of insect species is irrelevant; this role exists in the ecosystem, and so a species evolves to occupy this niche. A simple model of two species suggests that the impact niche tells us whether a coexistence point of two species is stable [?].

Both of these categories of semantics for the word niche have their use. There has been some work to resolve the discrepancies that arise when the two definitions conflict [10, 11]. This thesis tends to favour the requirement niche definition but ultimately remains agnostic to the debate. So long as niches exist in some sense, and a niche overlap parameter can be defined, the results I arrive at are sound. I felt it would be remiss were I not to include a brief summary of the debates associated with the definition of an ecological niche, hence the preceding section.

1.4.2 Lotka-Volterra

The original Lotka-Volterra model was introduced over a century ago to describe the dynamics of a population of a predator and its prey. It can be seen as an extension of the Verhulst, or logistic, equation, from one to two dimensions. These days the generalized Lotka-Volterra model is the accepted terminology for a dynamical system that depends linearly and quadratically on the populations modelled, with no explicit time dependence. A stochastic 2D model will be the main model used in this thesis, except for the next chapter, which exhaustively explores the stochastic Verhulst model. The deterministic limit of the 2D model has fixed points corresponding to neither species surviving, one, the other, or both. The position and stability of these points depends on the main parameters of the model, namely the growth rates, the carrying capacities, and the competition between species, called herein the niche overlap. Carrying capacity is a common phenomenological parameter that measures the number or density of organisms an ecosystem can support, in the absence of competitors. By growth rate I mean the timescale of approach toward the carrying capacity, typically measured experimentally by fitting a line to a semi-logarithmic plot of the growth curve. Some authors [5, 6, 12] have observed that for certain parameter values that the stochastic 2D generalized Lotka-Volterra model exhibits dynamics similar to those of the

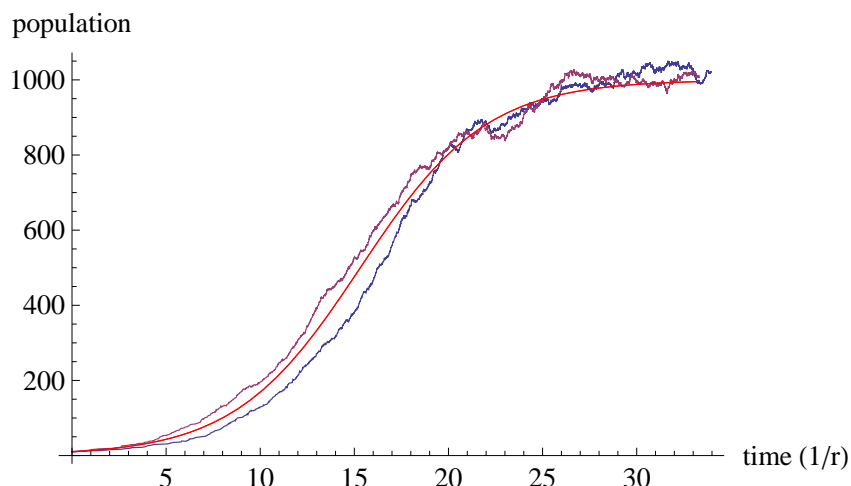


Figure 1.1: *A single logistic system with deterministic and stochastic solutions.* The smooth red line shows the deterministic solution to a one dimensional logistic differential equation with carrying capacity $K = 1000$, which the system asymptotically approaches. The jagged blue and purple lines are each an instantiation of a ‘noisy’, or stochastic, version of the logistic equation, as simulated using the Gillespie algorithm. Notice that the stochastic versions tend to follow their deterministic analogue but with some fluctuations, sometimes being greater than the deterministic result, sometimes being lesser.

Moran model. The transition to this limit is one of the main investigations of this thesis.

1.5 Stochastics

1.5.1 introduction

As stated before, a stochastic version of the two-dimensional generalized Lotka-Volterra model makes up the bulk of this thesis. What is meant by “stochastic”? Stochasticity is the technical term for randomness or noise in a system. Whereas the solution to a logistic differential equation would simply increase continuously (and differentiably) toward its asymptote at the carrying capacity, a stochastic version would allow for deviations from this trajectory, sometimes decreasing rather than steadily increasing toward the steady state, and thereupon fluctuating about the carrying capacity. See figure 1.5.1 for a visual example. In rare cases the fluctuations can even bring the system to a population of zero, in which case it does not recover. This is known as extinction, and is the main object of study in this thesis. Stochasticity has other uses too. It is the natural way to capture the difficulties of performing experiments, accounting for the imprecision of measurement and issues arising from sampling. More broadly, we need stochastics because of nature’s inherent randomness and because of the course-graining and phenomenological modelling necessarily done in biology (and indeed, in every scientific endeavor whose purview is not nanoscopic). There are applications in many disciplines, including linguistics, economics, biology, neuroscience, chemistry, and cryptography, to name a few. The giants Wright and Fisher were pioneers in applying randomness and statistical reasoning, in the biological context and in general. In biology there were renaissances in the stochastic treatment of genetics due to Kimura and ecology due to Hubbell.

1.5.2 Extinction rates from demographic and environmental stochasticity

It is a matter of common knowledge from the literature that demographic fluctuations lead to extinction times scaling exponentially in the system size, whereas environmental noise gives polynomial scaling [?]. That is, if K is the constant or mean system size, then demographic fluctuations lead to

$$\tau \propto e^{cK}$$

and environmental noise leads to

$$\tau \propto K^d,$$

for some constants c, d . This system size is often taken to be the carrying capacity [?]. This thesis only concerns itself with demographic fluctuations. Environmental fluctuations are a very real phenomenon that I nevertheless ignore. Consider this research as a null model; if the environment is constant then the results of the below research holds. Most real systems will not be represented by my results, but it gives a baseline against which to contrast.

The above extinction time scaling equations come from the Fokker-Planck equation. There are many ways to calculate the mean time to extinction (MTE). For most of my research I calculate the extinction time exactly, following a textbook formulation, or at least to arbitrary accuracy. There also exist many approximation techniques to deal with stochastic problems, as I will outline below.

1.5.3 Approximation techniques

The existence of a system size parameter K raises the possibility of approximation to the master equation, the equation which underlies all processes with demographic stochasticity. The aforementioned Fokker-Planck equation is an expansion of the master equation to continuous populations, going from a difference-differential equation to a partial differential equation. The results tend to look Gaussian about the deterministic dynamics and behave well near the fixed point. However, since extinction invariably happens near zero population, which is far from the fixed point for large system size, the Fokker-Planck approximation is expected to fail. It nevertheless does better than expected, and has utility in some contexts. It is also the easiest equation to use, both in terms of solution and further approximations, so it remains the most popular.

Recently popular is the WKB expansion. Rather than just expanding about the fixed point as is effectively the case for Fokker-Planck, WKB expands about the most probable trajectory. The WKB approach makes an ansatz solution to the master equation, which results in an effective Hamilton-Jacobi equation for some action-like object of the system. Upon solving the Hamiltonian mechanics the action need only be integrated along the route to fixation in order to estimate the mean time.

The main technique employed in this thesis is related to the formal solution to the master equation. In principle this involves inverting a semi-infinite matrix. By introducing a cutoff to the matrix I can calculate the mean fixation time. Varying the cutoff allows for arbitrary accuracy.

1.6 Structure of remaining thesis

The remaining structure of the thesis is as follows. First, I use the exact techniques introduced in section XXX to investigate a one dimensional logistic system, comparing the influence of the linear and quadratic

terms to the quasi-steady state distribution and the MTE. With the simplicity of this test system I explore the applicability of various common approximation techniques. The exact techniques and the approximations together make up chapter XXY, regarding a one dimensional system. This chapter is also being prepared as a paper for publication. The natural extension from a one dimensional logistic is to couple two such systems together; this arrives at the two dimensional generalized Lotka-Volterra system and is the subject of the next chapter, chapter XXZ. First a symmetric system is investigated, and the mean time to fixation is used as a tool to diagnose the longevity of the two interacting species. The overlap of their ecological niches is the parameter that controls the transition between effective coexistence and rapid fixation. Next the corresponding asymmetric model is explored. Along with the MTE, my analysis uncovers a typical route to fixation, the discussion of which wraps up this chapter. The final chapter introducing novel research, chapter XYX, extends the scope of this thesis to invasion of a new species into an already occupied niche. I calculate the probability of a successful invasion as a function of system size and niche overlap. Then the MTE conditioned on the success of the invasion is analyzed. Once these timescales are developed, I regard the Moran model modified to account for repeated invasions of the same species. Chapter XYY covers a variety of topics: I explore applications and extensions of the results arrived at in this thesis; I address the central problems introduced in this preliminary chapter and draw some conclusions informed by my results; and I suggest next steps for this research, both continuations and implementations to novel situations.

Chapter 2

Ch3-AsymmetricLogistic

2.1 Invasion Analysis

In the previous sections the two species obeyed symmetric dynamics, with random fluctuations leading to the eventual extinction of one or the other with equal likelihood. The mean time calculated was for this fixation, of either species. But if the symmetry is broken, for instance by starting away from the deterministic fixed point, one should not expect an equal likelihood of fixation and extinction for each species. Starting close to an axis leads to different timescales than those found above. For instance, in a scenario where a species is already settled in a niche, a small immigrant population enters or a small mutant strain arises, one that has some niche overlap with the established species. One might ask whether this invader will successfully establish itself, and if so, how long would a successful invasion take on average. Also of interest is the mean time of a failed invasion attempt. Both times set the scale against which one measures the immigration or mutation rate, to conclude whether a system should be a monoculture or show diversity. In any case we treat the situation where neither the established population nor the invader has an explicit fitness advantage. Each species has the same birth and death rates, equations ?? above with symmetric parameters. However, the system starts with $K - 1$ individuals of the established species and 1 invader. The invader strain is successful if it grows to be half of the total population before dying out. This success happens with probability E_s and in mean time τ_s . We also calculate the mean time conditioned on a failed invasion attempt, τ_f , for which the invading population never establishes itself.

As before, we expect $a = 1$, the complete niche overlap limit, to behave like the WFM model, and $a = 0$, the independent limit, to correspond to two single logistic systems. For $a = 1$ the Kramers-Moyal of the WFM result is [15]

$$E_s = 2/K, \tag{2.1}$$

$$\tau_s = \Delta t K^2 (K - 1) \ln \left(\frac{K}{K - 1} \right), \tag{2.2}$$

$$\tau_f = \Delta t (K - 2) \left(\ln(K) - (K - 1) \ln \left(\frac{K}{K - 1} \right) \right). \tag{2.3}$$

For $a = 0$ the invading mutant follows the dynamics of a single logistic system with carrying capacity K . Calculating E_s , τ_s , and τ_f for a one dimensional system is a textbook procedure [16] but does not have a compact solution and is shown in the Supporting Information. A one-dimensional deterministic logistic

system approaching its fixed point from $n = 1$ displays an invasion time that grows logarithmically with the system size (see Supplementary Information), and should also be a good match for τ_s .

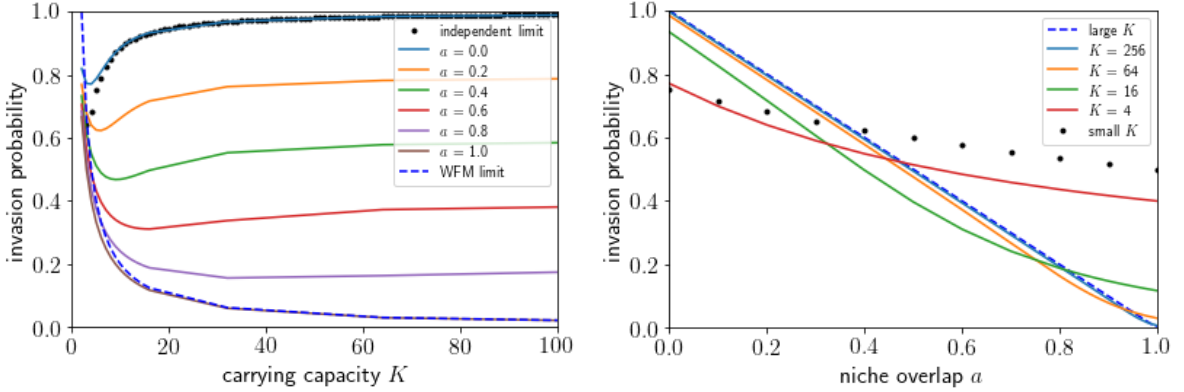


Figure 2.1: *Probability of a successful invasion.* Left: Solid lines are the numerical results, from $a = 0$ above to $a = 1$ below. The black dotted line is the expected single logistic limit, and the blue dashed line is WFM result. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), and matches well with the black dotted line $\frac{b_{mut}}{b_{mut} + d_{mut}}$ (see text for details). Successive lines are at larger system size, and approach the dashed blue line of $1 - (d_{mut}/b_{mut})$.

The calculated invasion probabilities and times are unintuitive, but regarding the asymptotic limits of small and large carrying capacity K allows for an understanding of the results. At small K commonly only a few birth or death events occur before invasion or extinction, and the slowest step determines the timescale. With $n_{mut} = 1$ and $n_{established} = K - 1$ this limiting step is the mutant death: with $K = 3$ the rates are $d_{mut} = (1 + 2a)/3$, $b_{mut} = 1$, $d_{est.} = (4 + 2a)/3 = 1 + d_{mut}$, $b_{est.} = 2$. Hence we expect $\tau \approx \frac{1}{d_{mut}} = \frac{K}{1 + a(K - 1)}$, with smaller niche overlap resulting in greater mean times, both of invasion and extinction. The invasion probability at this low K is the probability that the mutant reproduces before it dies, namely $\frac{b_{mut}}{b_{mut} + d_{mut}} = \frac{K}{K(1 + a) + 1 - a}$. In the other extreme, at large K , invasion will be likely and fast, as the stochastic drift draws the system to the deterministic fixed point for incomplete niche overlap. The invader birth and death rates can be approximated as constant for small invader number and large carrying capacity: a system with constant rates has an extinction probability of d_{mut}/b_{mut} [1, 2], which in this case implies $E_s \approx 1 - a$. The invasion probability go asymptotically to zero at large carrying capacity only in the WFM limit of $a = 1$.

Figures 2.1, 2.2, and 2.3 confirm our predictions for small and large K , respectively showing the invasion probability, mean time conditioned on invasion, and mean time conditioned on extinction of the invader. Note that at small carrying capacity the WFM limit has the shortest conditional times, but at large carrying capacity this complete niche overlap has the longest. In general, increasing a has conflicting effects; it brings the fixed point closer to the initial condition of one invader, suggesting a shorter timescale, but it also makes the two species more similar, effectively reducing the fast approach to the attractive deterministic fixed point. This non-monotonic dependence on a causes the unimodality of the conditional times in figures 2.2 and 2.3. An odd but reproducible feature, seen in figure 2.1, is that for some values of niche overlap there appears to be a minimum of probability at some intermediate carrying capacity. This is a low-number effect, and will not be relevant in most ecological systems, though in some situations it may apply, like with nascent cancer strains [3] or plasmid replacement [7]. Figure 2.3 also shows a maximum of mean time conditioned on a failed invasion attempt at intermediate

carrying capacity. It is consistent with the our expectation of fast times for both small and large K for incomplete niche overlap. The maximum appears for all values of a except $a = 1$.

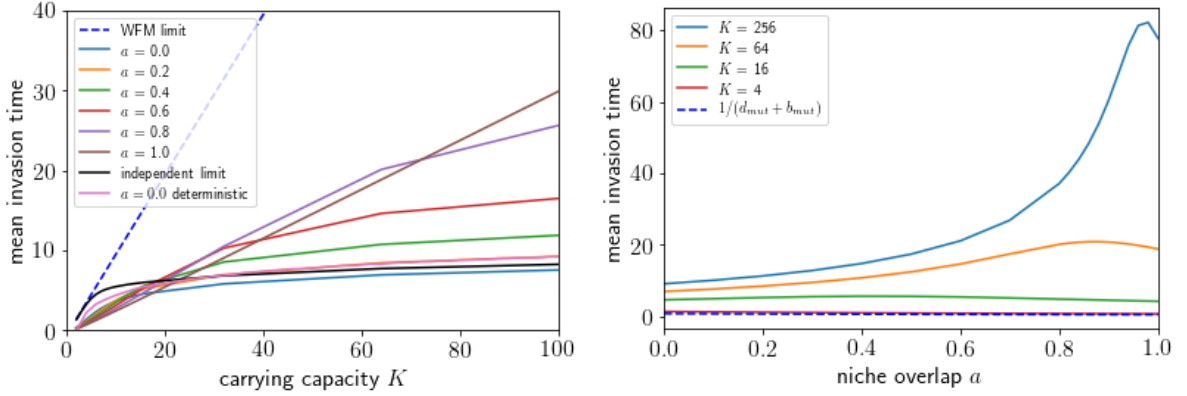


Figure 2.2: *Time of a successful invasion.* Left: Solid lines are the numerical results, from $a = 0$ at the bottom to $a = 1$ at top. The WFM result is given by the blue dashed line, and is linear, albeit with a slope that matches poorly with our results. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), and successive lines are at larger system size, up to $K = 256$. The dash dot black line is $1/d_{mut}$.

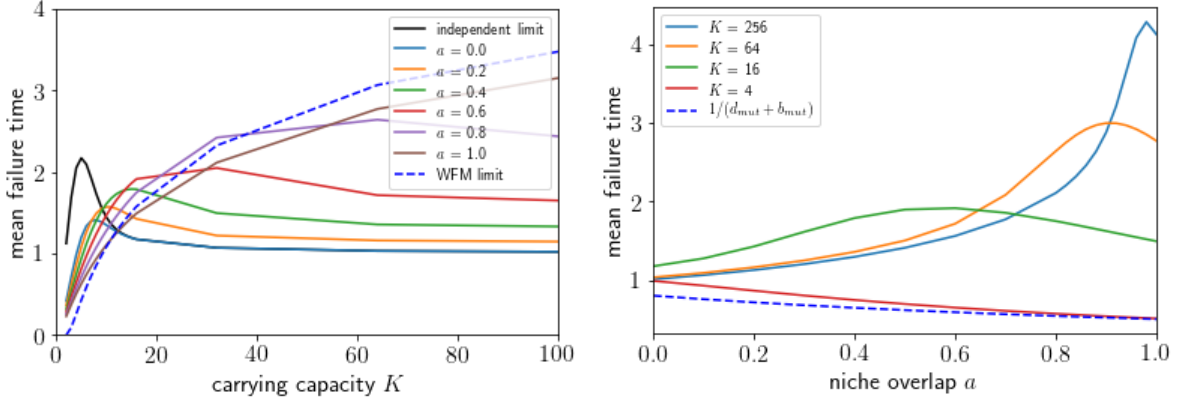


Figure 2.3: *Time of a failed invasion.* Left: Solid lines are the numerical results, from $a = 0$ mostly being fastest to $a = 1$ being slowest, for large K . The blue dashed line is WFM result. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), and successive lines are at larger system size, up to $K = 256$. The dash dot black line is $1/d_{mut}$.

These conditional times are hard to intuit. Increasing K moves the deterministic fixed point farther away, so we expect longer times, but it also draws the system more strongly, which would imply faster times. Regardless, all the timescales of invasion are fast: neither the mean time of a successful invasion nor of a failed attempt grows exponentially with the system size. This contrasts with the results of the previous sections, where the mean fixation time is exponential in the system size except when a is exactly one. τ_s at complete niche overlap has the fastest scaling, growing linearly with the carrying capacity. Whereas competing species will coexist for long times unless niche overlap is complete, the dynamics of the attempted invasion of a dominant species will be fast [?]. As will be discussed below, these timescales need to be compared against the mutation or immigration timescale to offer some insight on

mutation-selection balance.

2.2 Discussion

Unlike the fixation times of coexisting species in all but complete niche overlap, invasions into the system do not show exponential scaling in any limit. The likelihood of failure grows linearly with niche overlap, for sufficiently large K . For complete niche overlap the invasion probability goes asymptotically to zero, but it is low even for partially mismatched niches. The timescale of a successful invasion varies between linear and logarithmic in the system size. The mean time of an unsuccessful invasion is even faster than logarithmic, and for large K it becomes independent of K . Curiously, these failed invasion attempts are unimodal, at intermediate carrying capacity and niche overlap values. Our results provide a timescale to which the rate of immigration or mutation can be compared. If the influx of invaders is slower than the mean time of their failed invasion attempts, each attempt is independent and has the invasion probability we have calculated. However, if individual invaders arise faster than the time it takes to suppress the previous attempt, the new strains interact with one another in ways beyond the scope of this paper, leading to greater biodiversity. In the other extreme, if the time between invaders is even longer than the fixation times calculated in Section 4 then serial monocultures are expected.

2.3 Introduction

The Moran model [15] is a classic urn model used in population dynamics in a variety of ways. Its most prominent use is in coalescent theory, describing how the relative proportion of genes in a gene pool might change over time. But really it can describe any system where individuals of different species/strains undergo strong but unselective competition in some closed or finite ecosystem.

To arrive at the Moran model we must make some assumptions. Whether these are justified depends on the situation being regarded. The first assumption is that no individual is better than any other; that is, whether an individual reproduces or dies is independent of its species and the state of the system. This makes the Moran model a neutral theory, and any evolution of the system comes from chance rather than from selection.

Next we assume that the the population size is fixed, owing to the (assumed) strict competition in the system. That is, every time there is a birth the system becomes too crowded and a death follows immediately. Alternately, upon death there is a free space in the system that is filled by a subsequent birth. In the classic Moran model each pair of birth and death event occurs at a discrete time step (cf. the Wright-Fisher model, where each step involves N of these events). This assumption of discrete time can be relaxed without a qualitative change in results.

2.4 Moran Model in More Detail

In the classic Moran model, each iteration or time step involves a birth and a death event. Each organism is equally likely to be chosen (for either birth or death), hence a species is chosen according to its frequency, $f = n/N$, where N is the total population and n is the number of organisms of that species. Note that $N - n$ represents the remainder of the population, and need not all be the same species, so long as they are not the focal species denoted with ‘ n ’. The focal species increases in the population

if one of its members gives birth while a member of a different species dies; that is, $b(n) = f(1 - f)$. Increase and decrease of the focal species are equally likely, with

$$b(n) = f(1 - f) = (1 - f)f = d(n) = \frac{n}{N} \left(1 - \frac{n}{N}\right) = \frac{1}{N^2} n(N - n) \quad (2.4)$$

each time step Δt . Each step, the chance that nothing happens is $1 - (b(n) + d(n)) = f^2 + (1 - f)^2$. These are not rates themselves, rather they are the probability of an increase or decrease in the time step. A straightforward approximation would be to take Δt infinitesimal, then $b(n)\Delta t$ and $d(n)\Delta t$ serve as rates of birth and death of the species in a continuous time analogue to the Moran model.

For the record, here is the mean and variance as a function of time. If the system starts with n_0 individuals of the focal species, then there should be

$$(n_0 - 1)d(n_0) + (n_0 + 1)b(n_0) + n_0(1 - b(n_0) - d(n_0)) = n_0 - d(n_0) + b(n_0) = n_0$$

individuals in the next time step as well. Iterating this calculation gives that the expected value at all times is just the initial population, $\langle n \rangle(t) = n_0$. Given the delta function initial condition of starting with n_0 individuals, the variance should start at zero and grow. After one time step the second moment is

$$(n_0 - 1)^2 d(n_0) + (n_0 + 1)^2 b(n_0) + n_0^2 (1 - b(n_0) - d(n_0)) = n_0^2 - 2n_0 d(n_0) + 2n_0 b(n_0) + d(n_0) + b(n_0)$$

and the variance $V_1 = 2b(n_0) = 2d(n_0) = 2f_0(1 - f_0)$. For the variance at time step k we need the variance at $k - 1$ and the law of total variance, $E[Var(n_k | n_{k-1})] + Var(E[n_k | n_{k-1}]) = Var(n_k) \equiv V_k$. Recalling $E[n_k | n_{k-1}] = n_{k-1} = n_0$ and $Var(n_k | n_{k-1}) = 2f_{k-1}(1 - f_{k-1})$

$$\begin{aligned} V_k &= E[2f_{k-1}(1 - f_{k-1})] + Var(n_{k-1}) \\ &= 2\langle f_{k-1} \rangle - 2\langle n_{k-1}^2 \rangle / N^2 + V_{k-1} \\ &= 2\langle f_{k-1} \rangle - 2(V_{k-1} + \langle n_{k-1} \rangle^2) / N^2 + V_{k-1} \\ &= 2\langle f_{k-1} \rangle (1 - \langle f_{k-1} \rangle) + (1 - 2/N^2) V_{k-1} \\ &= V_1 + (1 - 2/N^2) V_{k-1}. \end{aligned}$$

Iterating the above and using the geometric series $\sum_{i=0}^{k-1} r^i = (1 - r^k)/(1 - r)$ gives

$$V_k = V_1 (1 - (1 - 2/N^2)^k) / (2/N^2) = n_0(N - n_0) (1 - (1 - 2/N^2)^k).$$

Notice that as $N \rightarrow \infty$ the variance, a measure of the fluctuations, goes to zero, and the system becomes deterministic. [maybe cf. hardy-weinberg variances] For finite N the variance goes to $N^2 f_0(1 - f_0)$ at long times.

The system fluctuates until either the species dies (extinction) or all others die (fixation). Both of these cases are absorbing states, so once the system reaches either it will never change. Since a species is equally likely to increase or decrease each time step, the model is akin to an unbiased random walk, and therefore the probability of extinction occurring before fixation is just

$$E(n) = 1 - n/N = 1 - f. \quad (2.5)$$

DERIVE THIS!!!!!! The first passage time, however, does not match a random walk, as there is a probability of no change in a time step, and this probability varies with f . DERIVE THE FIRST PASSAGE TIMES AS WELL (conditional and un?!?!)

We define the unconditioned first passage time $\tau(n)$ as the time the system takes, starting from n organisms of the focal species, to reach either fixation *or* extinction. It can be calculated by regarding how the mean from one starting position n relates to the mean of its neighbours.

$$\tau(n) = \Delta t + d(n)\tau(n-1) + (1-b(n)-d(n))\tau(n) + b(n)\tau(n+1) \quad (2.6)$$

Subbing in the values of the ‘birth’ and ‘death’ rates and rearranging this gives

$$\tau(n+1) - 2\tau(n) + \tau(n-1) = -\frac{\Delta t}{b(n)} = -\Delta t \frac{N^2}{n(N-n)}, \quad (2.7)$$

or

$$\tau(f+1/N) - 2\tau(f) + \tau(f-1/N) = -\Delta t \frac{1}{f(1-f)}. \quad (2.8)$$

If we approximate the LHS of the above with a double derivative (ie. $1 \ll N$) we get

$$\frac{\partial^2 \tau}{\partial n^2} = -\Delta t N \left(\frac{1}{n} + \frac{1}{N-n} \right) \quad (2.9)$$

Double integrate and use the bounds $\tau(0) = 0 = \tau(N)$ to get

$$\tau(n) = -\Delta t N^2 \left(\frac{n}{N} \ln \left(\frac{n}{N} \right) + \frac{N-n}{N} \ln \left(\frac{N-n}{N} \right) \right). \quad (2.10)$$

Note that we didn’t need to use the large N approximation: there is an exact solution:

$$\tau(n) = \Delta t N \left(\sum_{j=1}^n \frac{N-n}{N-j} + \sum_{j=n+1}^N \frac{n}{j} \right). \quad (2.11)$$

2.5 Moran With Immigration

Previous sections have stated that different dynamics are expected depending on a comparison of timescales. If new species enter the system faster than they go extinct, the biodiversity should increase to some steady state. Conversely, if extinction is much more rapid than speciation, a monoculture is expected in the system. Whether the monocultural system contains the same species over multiple invasion attempts or whether it experiences sweeps, changing from a monoculture of one species to the next, depends on the probability of a successful invasion. To arrive at some analytic solutions, we will treat a simplified model.

The basis of the following model is that of Moran, with its finite population size and discrete time steps, although we will relax the latter constraint. For comparison, Crow and Kimura [? ?] treat the problem with both continuous time and continuous populations (ie. population densities), arriving at some numerical results but not much else... Our inspiration is an /interesting/ work from the Gore lab [19], measuring the gut microbiome of bacteria-consuming *C. elegans* grown in a 50-50 environment of two strains of fluorescently-labeled but otherwise identical *E. coli*. After an initial colonization period, each

nematode has a stable number of bacteria in their gut, presumably from a balance of immigration, birth, and death/emigration. The researchers find a distribution of populations depending on the comparison of two experimental timescales.

For the model in this section, consider a focal species of n organisms, with the remaining $N - n$ organisms being of a different strain (or strains). Again we define a fractional abundance $f = n/N$. Traditionally the Moran population is thought to be some isolated population, and immigrants come from some metapopulation of larger size and diversity. We shall see if the Moran population acts as a reservoir, and generally what its dynamics are. The metapopulation has the same species we were originally talking about, with m , M and g analogous to n , N and f . That is, assume the immigrant into the Moran population is a member of the focal species with probability g , and not that species with probability $1 - g$. In theory g should be a random number drawn from the probability distribution associated with some evolving metapopulation, but for now we will take it to be fixed. That is, we assume that the metapopulation changes much slower than the Moran population of interest. In the analogy of the Gore experiment, the system of interest is the nematode gut, and the metapopulation is the environment in which the nematode lives (and eats). The consumption of one bacterium will influence the gut microbiome while having a negligible effect on the external environment.

Suppose immigration acts like birth in the Moran model. That is, ν of the time an immigrant comes in instead of a birth event occurring. Death occurs as normal. Then we have the following possibilities:

transition	function	value
$n \rightarrow n + 1$	$b(n)$	$f(1 - f)(1 - \nu) + \nu g(1 - f)$
$n \rightarrow n - 1$	$d(n)$	$f(1 - f)(1 - \nu) + \nu(1 - g)f$
$n \rightarrow n$	$1 - b(n) - d(n)$	$(f^2 + (1 - f)^2)(1 - \nu) + \nu(gf + (1 - g)(1 - f))$

Note that the birth and death rates are no longer the same as each other (as they are, in the classical Moran model); there is a bias in the system, toward g . Just as with the classical Moran model, strictly speaking b and d are probabilities rather than rates. The continuous time model, which well approximates the discrete time Moran, is attained by calling b and d rates and taking Δt to zero.

If a new mutant or immigrant species is unlikely to enter again (ie. if $g \simeq 0$) then this is close to the regular Moran model, and will not be treated further here. The system then corresponds to the regular Moran model presented in the introduction. Here we regard the case where it is possible to draw in the species of interest from the metacommunity, before it goes extinct in the focus community (ie. $\nu g \gg 1/\tau$). Since there will be always be immigration, the system will never truly fixate, as there will always be immigrants of the ‘extinct’ species to be reintroduced to the population. Rather, the system will settle on a stationary distribution. The process will have the master equation $\frac{dP_n(t)}{dt} = P_{n-1}(t)b(n-1) + P_{n+1}(t)d(n+1) - (b(n) + d(n))P_n(t)$, which gives a difference relation when the time derivative is set to zero. Since the system is constrained between 0 and N we normalize the finite number of probabilities and sum them to unity to get

$$\tilde{P}_n = \frac{q_n}{\sum_{i=0}^{\infty} q_i} \quad (2.12)$$

where

$$\begin{aligned} q_0 &= \frac{1}{b(0)} = \frac{1}{\nu g} \\ q_1 &= \frac{1}{d(1)} = \frac{N^2}{(N-1)(1-\nu) + \nu N(1-g)} \\ q_i &= \frac{b(i-1) \cdots b(1)}{d(i)d(i-1) \cdots d(1)}, \quad i \geq 1 \\ &= \prod_{j=1}^{i-1} \frac{1}{d(j)} \sum_{j=1}^{i-1} \frac{b(j)}{d(j)} \end{aligned}$$

recalling that $\frac{b(i)}{d(i)} = \frac{i(N-i)(1-\nu) + \nu N g(N-i)}{i(N-i)(1-\nu) + \nu N(1-g)i}$. The unnormalized steady-state probability can be written compactly as

$$\begin{aligned} q_n &= \frac{N^2 \Gamma(N) \Gamma\left(n + \frac{gN\nu}{1-\nu}\right) \Gamma\left(N - n + 1 + \frac{(1-g)N\nu}{1-\nu}\right)}{(n(N-n)(1-\nu) + (1-g)nN\nu) \Gamma(n) \Gamma(N-n+1) \Gamma\left(1 + \frac{gN\nu}{1-\nu}\right) \Gamma\left(N + \frac{(1-g)N\nu}{1-\nu}\right)} \\ q_n &= \frac{N^2 (N-1)! \left(n - 1 + \frac{gN\nu}{1-\nu}\right)! \left(N - n + \frac{(1-g)N\nu}{1-\nu}\right)!}{\left(n(N-n)(1-\nu) + (1-g)nN\nu\right) (n-1)! (N-n)! \left(\frac{gN\nu}{1-\nu}\right)! \left(N - 1 + \frac{(1-g)N\nu}{1-\nu}\right)!} \end{aligned}$$

and the sum of these is the normalization $\sum q_i = {}_2F_1(-N, gN\nu/(1-\nu); 1 - N(1-g\nu)/(1-\nu); 1)/g\nu$ which follows from the hypergeometric definition and q_i

$$\sum q_i = \frac{1}{g\nu} \frac{\Gamma[1 - \frac{N(1-g\nu)}{1-\nu}] \Gamma[N + 1 - \frac{N}{1-\nu}]}{\Gamma[N + 1 - \frac{N(1-g\nu)}{1-\nu}] \Gamma[1 - \frac{N}{1-\nu}]} = \frac{1}{g\nu} \frac{(-\frac{N(1-g\nu)}{1-\nu})! (-\frac{N\nu}{1-\nu})!}{(-\frac{N(1-g)\nu}{1-\nu})! (-\frac{N}{1-\nu})!}$$

which together gives \tilde{P}_n . But I should be careful, because I think I summed this to infinity, rather than to N - checked; it makes no difference apparently (and anyway assume $q_{n>N} = 0$).

$$Pochhammer[a, n] = (a)_n = \Gamma(a+n)/\Gamma(a)$$

$$\Gamma(n) = (n-1)! = \int_0^\infty t^{n-1} e^{-t} dt$$

$\ln(-x) = \ln(x) + i\pi$ [yes] for $x > 0$ and $\Gamma(-x) = (-(x+1))! = (x+1)! + i\pi = ?\Gamma(x+2)?$ [no] - I'm not sold that this line is true!!!

Stirling: $\ln n! \approx n \ln n - n$ so $\ln \Gamma(n) = \ln n!/n \approx n \ln n - 2n$

$$Hypergeometric2F1[a, b; c; z] = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 \frac{t^{b-1} (1-t)^{c-b-1}}{(1-tz)^a} dt = \sum_{n=0}^\infty \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!} = (1-z)^{c-a-b} {}_2F_1(c-a-b, c-b; c; z)$$

$${}_2F_1(a, b; c; 1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$$

Since $q_1 = 1$ the stationary probability at 1 is \tilde{P}_1 ; this gives the flux to 0, hence the exit times. Similarly $n = N-1$ should be the other place whence it exits (but it's not clear whether $q_{N-1} = 1$). See figure 2.4 for a visualization of the steady-state probability distribution for different immigration/speciation rates.

We can easily calculate the mean and variance as a function of time before reaching steady state. If the mean μ at some time step k has $\mu_k = n_k$ individuals, then after one time step there should be $\mu_{k+1} = n_k - d(n_k) + b(n_k) = n_k + \nu(g - f_k)$ individuals. That is, $\mu_{k+1} - \mu_k = \nu(g - \mu_k/N)$. This is

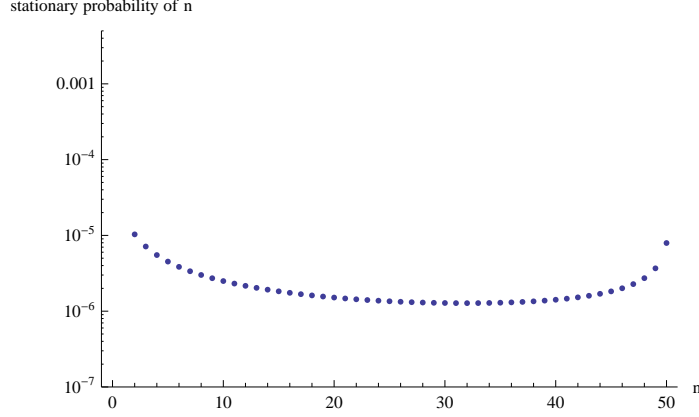


Figure 2.4: PDF of stationary Moran process due to immigration. $g = 0.1$, $N = 50$, $\nu = 0.01$.

solved by

$$\mu_k = \langle n \rangle(k) = gN \left(1 - (1 - n_0)(1 - \nu/N)^k \right).$$

At long times the mean fraction f matches that of the metapopulation, g . To get the an approximation of the variance, we will consider the continuous time analogue. First, the mean evolves as $\partial_t \mu = \langle b(n) - d(n) \rangle = \nu(g - \mu/N)$, which has the solution $\mu(t) = gN + (\mu_0 - gN)e^{-\nu t/N}$, and the timescale is set by ν/N . The dynamical equation for the second moment is

$$\begin{aligned} \partial_t \langle n^2 \rangle &= 2\langle nb(n) - nd(n) \rangle + \langle b(n) + d(n) \rangle \\ &= 2\nu(g\mu - \langle n^2 \rangle/N) + 2(1 - \nu)(N\mu - \langle n^2 \rangle)/N^2 + \nu(\mu + gN - 2\mu g)/N \end{aligned}$$

which is an inhomogeneous linear differential equation. The solution is long but not complicated. Recalling that $\sigma^2(t) = \partial_t \langle n^2 \rangle(t) - \mu^2(t)$ I write the variance as

$$\sigma^2(t) = \sigma^2(\infty) + (1 + g\nu - g(1 - \nu)/N)N^2 \frac{\mu_0 - gN}{N\nu + 2(1 - \nu)} \exp\left\{-\frac{\nu}{N}t\right\} - (gN - \mu_0)^2 \exp\left\{-2\frac{\nu}{N}t\right\} + C \exp\left\{-\frac{2}{N}\left(\nu + \frac{(1 - \nu)}{N}\right)t\right\}$$

where C is an integration constant, and is $C = \sigma^2(0) - \sigma^2(\infty) + (gN - \mu_0)^2 + (gN - \mu_0)(2 - \nu)(1 - 2g)/(N\nu + 2(1 - \nu))$ if the initial variance is $\sigma^2(0)$. $\sigma^2(\infty) = g(1 - g)N^2 \frac{1}{1 + \nu(N - 1)}$ is the long time, steady state variance of the system.

Figure 2.4 gives the probability distribution of the species of interest averaged over long times, but does not allow us to infer anything about the time scales or dynamics of the system. To do so, we must look at a slightly modified problem, with modified transition rates such that $b(0) = d(N) = 0$. This allows us to find the mean first passage time to species fixation or extinction, recognizing that this will only be a temporary state. The technique follows that laid out in the introduction. As a brief reminder, define E_i as the probability that the focal species goes extinct in this modified system with absorbing states at $n = 0$ and $n = N$, ie. the system goes to the former before the latter, given that it starts at $n = i$. Then $E_i = \frac{b(i)}{b(i) + d(i)} E_{i+1} + \frac{d(i)}{b(i) + d(i)} E_{i-1}$. Further define $S_i = \frac{d(i) \cdots d(1)}{b(i) \cdots b(1)}$. Then

$$E_i = \frac{\sum_{j=i}^{N-1} S_j}{1 + \sum_{j=1}^{N-1} S_j}. \quad (2.13)$$

See figure 2.5 for a graphical representation of the results. As with the stationary distribution, the extinction probabilities can be written explicitly, but the solution has an even less nice form.

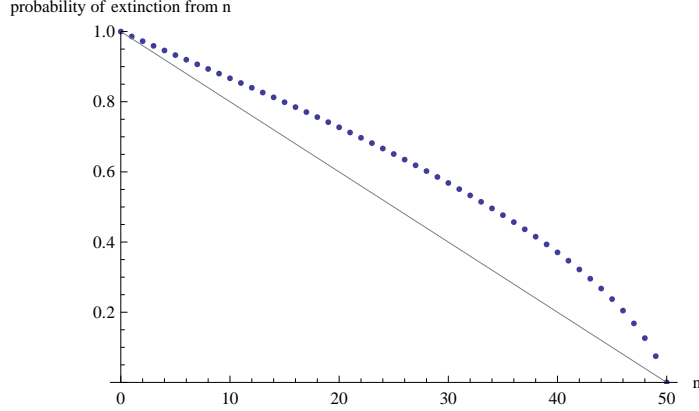


Figure 2.5: Probability of first going extinct, given starting population/fraction. $g = 0.1$, $N = 50$, $\nu = 0.01$. Grey is regular Moran results without immigration.

Similar to the extinction probabilities, we can write unconditioned mean first passage times to get

$$\tau[i] = \sum_{k=1}^{N-1} q_k + \sum_{j=1}^{i-1} S_j \sum_{k=j+1}^{N-1} q_k \quad (2.14)$$

where

$$q_i = \frac{b(i-1) \cdots b(1)}{d(i)d(i-1) \cdots d(1)}.$$

$$\text{so ultimately } \tau[n] = -\frac{N^2}{-u+N(gu-1)+1} + \sum_{j=2}^{n-1} \frac{\Gamma(j+1) \left(\frac{-guN+N+u-1}{u-1} \right)_j \left(\frac{g(-u+N(gu-1)+1)(1-N)_{N-1} \left(1 - \frac{gNu}{u-1} \right)_{N-1} + (g-1)\Gamma(N) \left(guN^2 - guN + N + u + (-u+N(gu-1)+1) {}_2F_1 \left(-N, -\frac{gNu}{u-1}; \frac{-guN+N+u-1}{u-1}; 1 \right) - 1 \right) \left(\frac{-guN+N+u-1}{u-1} \right)_{N-1}}{(g-1)gu(-u+N(gu-1)+1)\Gamma(N) \left(\frac{-guN+N+u-1}{u-1} \right)_{N-1}}}{(g-1)g(-u+N(gu-1)+1)(1-N)_{N-1} \left(1 - \frac{gNu}{u-1} \right)_{N-1} + (g-1)\Gamma(N) \left(guN^2 - guN + N + u + (-u+N(gu-1)+1) {}_2F_1 \left(-N, -\frac{gNu}{u-1}; \frac{-guN+N+u-1}{u-1}; 1 \right) - 1 \right) \left(\frac{-guN+N+u-1}{u-1} \right)_{N-1}}$$

Note that this should go to zero at both $n = 0$ and $n = N$, since it is unconditioned. Again, there is a close form, but it is a sum of hyperbolic functions and has not been fully explored yet. It is approximated numerically and displayed graphically in figure 2.6.

Keeping with the artificial stoppage when the focal population reaches 0 or N individuals, we calculate the conditional times, respectively to extinction and to fixation. As before, the extinction probability is given by equation 2.13. Similar to the continuous time solutions presented in the introduction, the conditional extinction time can be written as

$$\phi[n] = \phi[1] + \sum_{j=1}^{n-1} \left(\phi[1] - \sum_{i=1}^j q_i E_i \right) S_j. \quad (2.15)$$

We have the other boundary condition that, since $\theta_N = 0$, $\phi_N = 0$, which allows us to rearrange the previous equation to get

$$\phi_1 = \frac{\sum_{j=1}^{N-1} \sum_{i=1}^j q_i E_i}{1 + \sum_{j=1}^{N-1} S_j}. \quad (2.16)$$

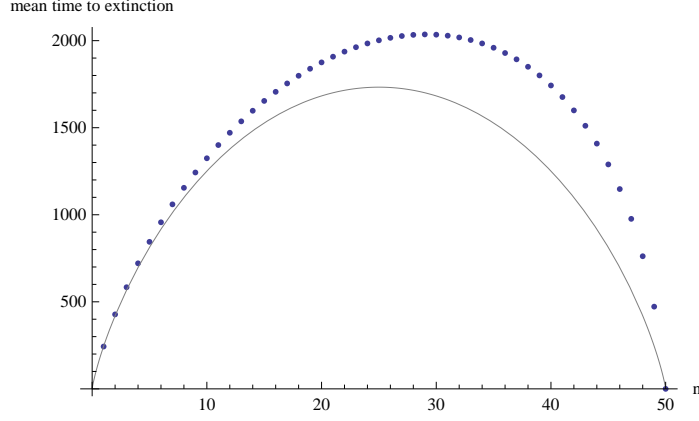


Figure 2.6: Mean time to either fixation or extinction, given starting population/fraction. $g = 0.1$, $N = 50$, $\nu = 0.01$. Grey is regular Moran results without immigration.

Here $\phi_i \equiv E_i \theta_i$ (not a dot product, just multiplication of elements), where θ_i is the conditional extinction time. These previous two equation allow us to solve for ϕ , and therefore θ . After all this, one arrives at the graph in figure 2.7. The conditional times mostly follow the unconditioned time, except near the rare events that do not much contribute to the average.

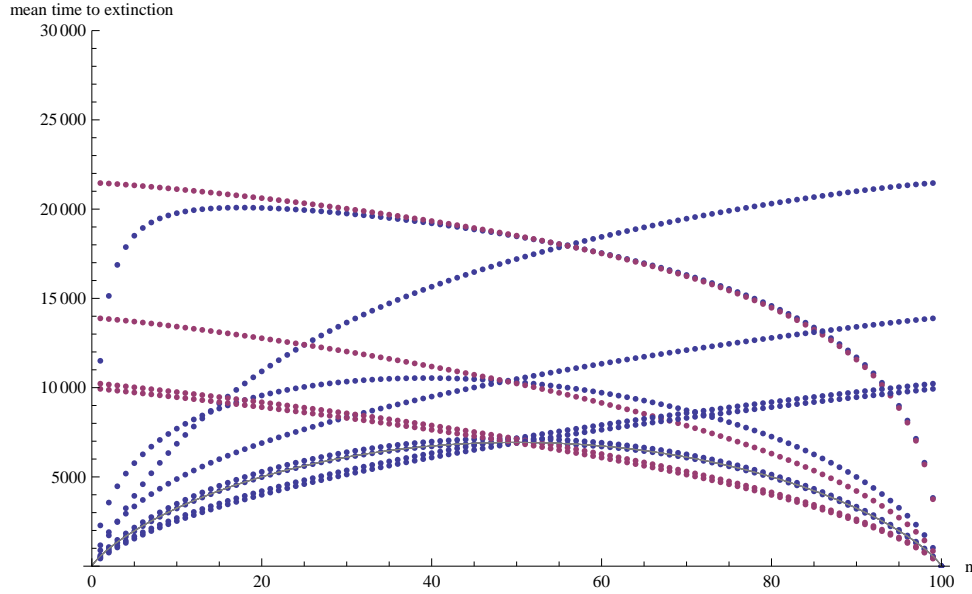


Figure 2.7: Mean time to fixation or extinction, conditioned on that event happening, given starting population/fraction. $g = 0.7$, $N = 100$, ν varies from 0.3 (highest) to 0.0001 (lowest). Grey is regular Moran results without immigration.

2.6 Some Results

With all that we've discovered, we can say something about the switching behaviour of this Moran population with immigrants. Likely, there are some limiting forms of the analytic expressions that will offer more insight - these are currently being investigated. Let's say the population starts in state $n = 0$.

Then at a rate νg there will be an attempted invasion. The invasion will be successful only every $1/E_1$ attempts. Thus a successful invasion occurs every $1/\nu g E_1$ time units. We can compare the time between successful invasions, and the time between attempted invasions, with the time each attempt takes (successful or not). COMPARE SOME ACTUAL NUMBERS ALSO REFER BACK TO THE GRAPHS

Bibliography

- [1] Edward J Allen, Linda J S Allen, and Henri Schurz. *A comparison of persistence-time estimation for discrete and continuous stochastic population models that include demographic and environmental variability.*, volume 196. jul 2005.
- [2] Linda J S Allen. *An Introduction to Stochastic Processes with Applications to Biology.* Pearson Education, Inc, New Jersey, 2003.
- [3] Peter Ashcroft, Franziska Michor, and Tobias Galla. Stochastic tunneling and metastable states during the somatic evolution of cancer. *Genetics*, 199(4):1213–1228, 2015.
- [4] Peter Chesson. MECHANISMS OF MAINTENANCE OF SPECIES DIVERSITY. *Annu. Rev. Ecol. Syst.*, 31(May):343–66, 2000.
- [5] Thiparat Chotibut and David R. Nelson. Evolutionary dynamics with fluctuating population sizes and strong mutualism. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.*, 92(2), 2015.
- [6] George W.A. Constable and Alan J. McKane. Models of genetic drift as limiting forms of the Lotka-Volterra competition model. *Phys. Rev. Lett.*, 114(3):1–5, 2015.
- [7] Robert Gooding-townsend, Steven T E N Holder, and Brian Ingalls. Displacement of Bacterial Plasmids by Engineered Unilateral Incompatibility. (April):19–21, 2015.
- [8] S.P. Hubbell. *The Unified Theory of Biodiversity and Biogeography.* 2001.
- [9] G.E. Hutchinson. The Paradox of the Plankton. 95(882):137–145, 1961.
- [10] M A Leibold and Mark A McPeck. Coexistence of the Niche and Neutral Perspectives in Community Ecology. *Ecology*, 87(6):1399–1410, 2006.
- [11] Matthew A Leibold. The Niche Concept Revisited : Mechanistic Models and Community Context. *Ecology*, 76(5):1371–1382, 1995.
- [12] Yen Ting Lin, Hyejin Kim, and Charles R. Doering. Features of Fast Living: On the Weak Selection for Longevity in Degenerate Birth-Death Processes. *J. Stat. Phys.*, 148(4):646–662, 2012.
- [13] Robert Macarthur and Richard Levins. Divergence of Coexisting Species. 101(921):377–385, 1967.
- [14] R May. Unanswered questions in ecology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 354(1392):1951–9, dec 1999.
- [15] Patrick Moran. *The Statistical Processes of Evolutionary Theory.* Clarendon Press, Oxford, 1962.

- [16] R.M. Nisbet and W.S.C. Gurney. *Modelling Fluctuating Populations*. John Wiley & Sons, Toronto, 1982.
- [17] Elizabeth Pennisi. What determines species diversity? *Science* (80-.), 309:90, 2005.
- [18] Shovonlal Roy and J. Chattopadhyay. Towards a resolution of 'the paradox of the plankton': A brief overview of the proposed mechanisms. *Ecol. Complex.*, 4(1-2):26–33, 2007.
- [19] Nicole M. Vega and Jeff Gore. Stochastic assembly produces heterogeneous communities in the *Caenorhabditis elegans* intestine. *PLoS Biol.*, 15(3):1–20, 2017.