

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	Ch3-AsymmetricLogistic	1
1.1	Invasion Analysis	1
1.2	Discussion	4
1.3	Introduction	4
1.4	Moran Model in More Detail	4
1.5	Moran With Immigration	6
1.6	Some Results	12
	Bibliography	13

Chapter 1

Ch3-AsymmetricLogistic

1.1 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]

$$E_s = 2/K, \tag{1.1}$$

$$\tau_s = \Delta t K^2 (K - 1) \ln \left(\frac{K}{K - 1} \right), \tag{1.2}$$

$$\tau_f = \Delta t (K - 2) \left(\ln(K) - (K - 1) \ln \left(\frac{K}{K - 1} \right) \right). \tag{1.3}$$

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 .

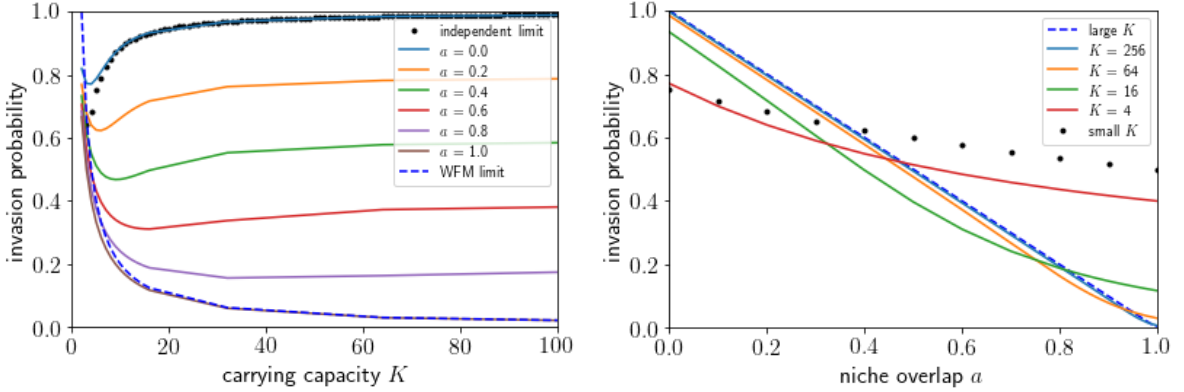


Figure 1.1: *Probability of a successful invasion.* Left: Solid lines are the numerical results, from $a = 0$ above to $a = 1$ below. The black dotted line is the expected single logistic limit, and the blue dashed line is WFM result. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), 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 1.1, 1.2, and 1.3 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 1.2 and 1.3. An odd but reproducible feature, seen in figure 1.1, 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 1.3 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 $a = 1$.

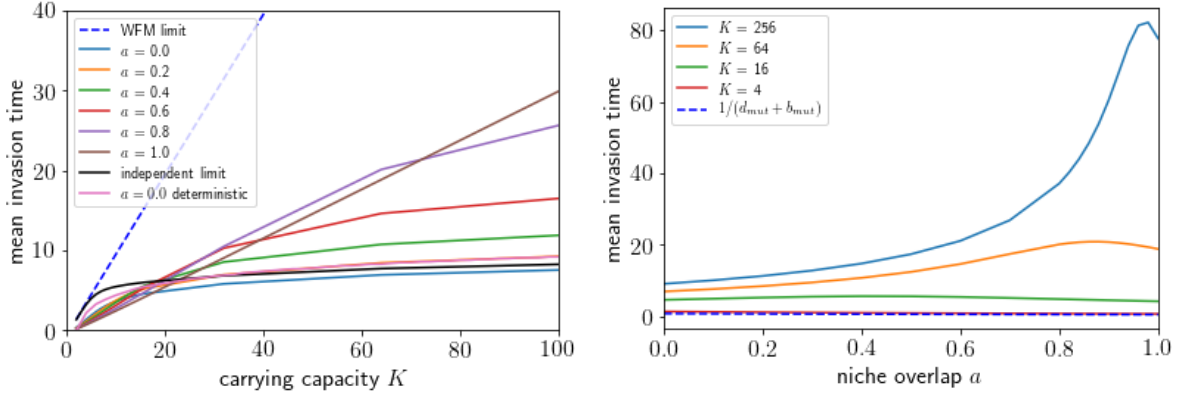


Figure 1.2: *Time of a successful invasion.* Left: Solid lines are the numerical results, from $a = 0$ at the bottom to $a = 1$ at top. The WFM result is given by the blue dashed line, and is linear, albeit with a slope that matches poorly with our results. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), and successive lines are at larger system size, up to $K = 256$. The dash dot black line is $1/d_{mut}$.

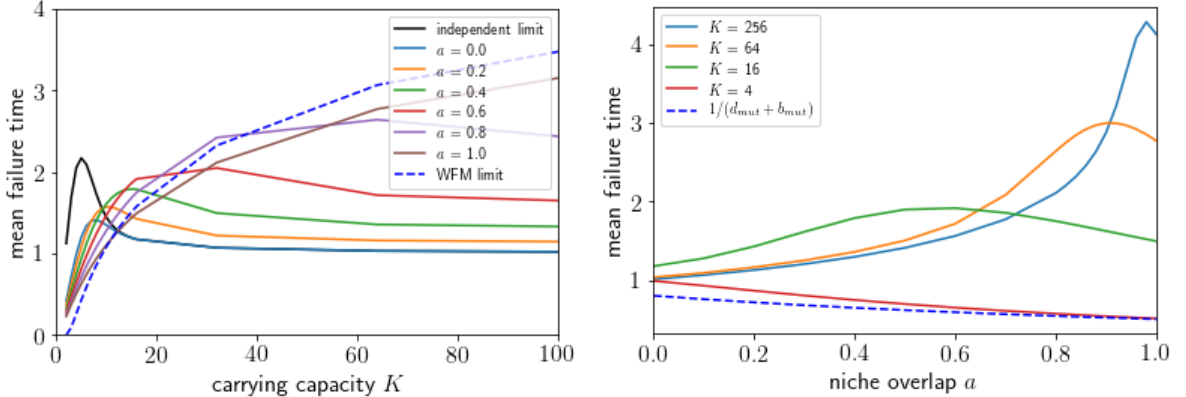


Figure 1.3: *Time of a failed invasion.* Left: Solid lines are the numerical results, from $a = 0$ mostly being fastest to $a = 1$ being slowest, for large K . The blue dashed line is WFM result. Right: The solid blue line shows the results for small carrying capacity ($K = 4$), and successive lines 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.

1.2 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.

1.3 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.

1.4 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

$$b(n) = f(1 - f) = (1 - f)f = d(n) = \frac{n}{N} \left(1 - \frac{n}{N}\right) = \frac{1}{N^2} n(N - n) \quad (1.4)$$

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. This corresponds to f_0 of the probability mass being at $n = N$, and $(1 - f_0)$ being at $n = 0$, since at long times the system has fixated at one end or the other.

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

$$E(n) = 1 - n/N = 1 - f. \quad (1.5)$$

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.

$$\tau(n) = \Delta t + d(n)\tau(n-1) + (1-b(n)-d(n))\tau(n) + b(n)\tau(n+1) \quad (1.6)$$

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

$$\tau(n+1) - 2\tau(n) + \tau(n-1) = -\frac{\Delta t}{b(n)} = -\Delta t \frac{N^2}{n(N-n)}, \quad (1.7)$$

or

$$\tau(f+1/N) - 2\tau(f) + \tau(f-1/N) = -\Delta t \frac{1}{f(1-f)}. \quad (1.8)$$

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

$$\frac{\partial^2 \tau}{\partial n^2} = -\Delta t N \left(\frac{1}{n} + \frac{1}{N-n} \right) \quad (1.9)$$

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

$$\tau(n) = -\Delta t N^2 \left(\frac{n}{N} \ln \left(\frac{n}{N} \right) + \frac{N-n}{N} \ln \left(\frac{N-n}{N} \right) \right). \quad (1.10)$$

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

$$\tau(n) = \Delta t N \left(\sum_{j=1}^n \frac{N-n}{N-j} + \sum_{j=n+1}^N \frac{n}{j} \right). \quad (1.11)$$

1.5 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

$$\tilde{P}_n = \frac{q_n}{\sum_{i=0}^{\infty} q_i} \quad (1.12)$$

where

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

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

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

and the sum of these is the normalization

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

which follows formally from the definition of the hypergeometric function ${}_2F_1$. Together these give \tilde{P}_n . See figure 1.4 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 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

$$\sigma^2(t) = \sigma^2(\infty) + A \exp\left\{-\frac{\nu}{N}t\right\} - B \exp\left\{-2\frac{\nu}{N}t\right\} + C \exp\left\{-\frac{2}{N}\left(\nu + \frac{(1-\nu)}{N}\right)t\right\}$$

where $A = (1 + g\nu - g(1-\nu)/N)N^2 \frac{\mu_0 - gN}{N\nu + 2(1-\nu)}$, $B = (gN - \mu_0)^2$, and C is an integration constant;

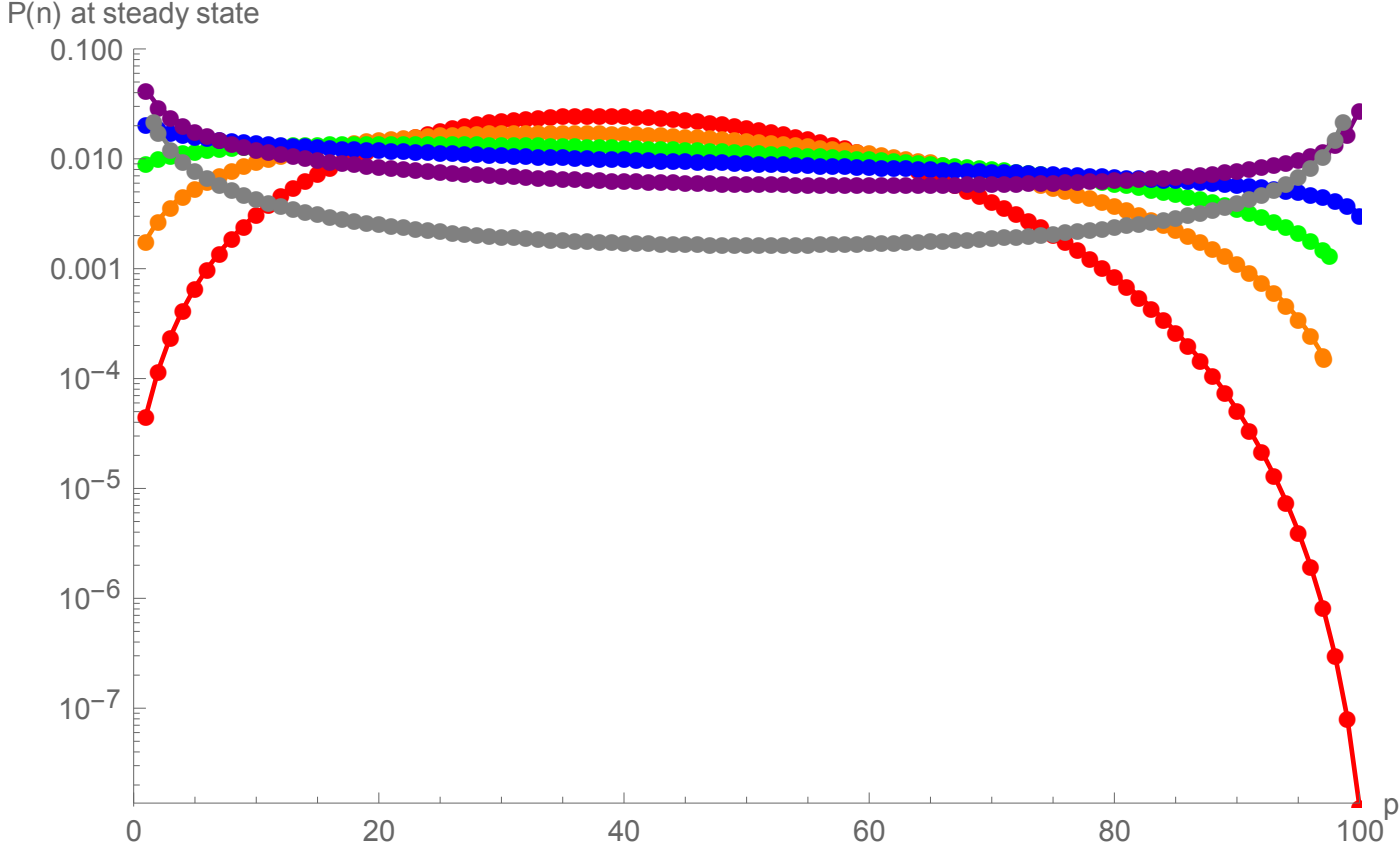


Figure 1.4: PDF of stationary Moran process due to immigration. $g = 0.4$, $N = 100$, $\nu * N$ is given by the colour; red is 10, orange is 5, green is 3, blue is 2, purple is 1, and grey is 0.2. Notice that the curvature of the distribution inverts around $\nu = 2/N$.

$C = \sigma^2(0) - \sigma^2(\infty) + (gN - \mu_0)^2 + (gN - \mu_0)(2 - \nu)(1 - 2g)/(N\nu + 2(1 - \nu))$ 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.

Notice that for $g = 0, 1$ the long term variance $\sigma^2(\infty)$ goes to zero. This contrasts with the results of the Moran model without immigration, where a fraction of instances fixate with the focal species and in the remaining fraction that species goes extinct, in proportion to its initial abundance. Having a supply of immigrants destabilizes one of these absorbing states, such that the ultimate fate is either none of the focal species for $g = 0$ or only the focal species for $g = 1$. The memory of the initial abundance does not affect these results at long times. However, if the immigration rate is truly small, such that $N\nu \ll 1$, we recover similar results to the no immigration case. Instead of $f_0(1 - f_0)N^2$ we get $\sigma^2(\infty) \approx g(1 - g)N^2$, with the metapopulation focal species abundance g acting as the initial abundance. This is because the fixation time of the Moran model, which goes like N , is much faster than the immigration time $1/\nu$. Upon entry of a new immigrant the Moran model fixates as usual, in proportion to either $1/N$ or $(N - 1)/N$, depending on the species of the immigrant, which in turn is governed by the metapopulation abundance g . Each iteration goes one way or the other, typically to the closest extreme, which a fraction g of the time is the focal species, hence $\sigma^2(\infty) \approx g(1 - g)N^2$. The fixation need not happen more rapidly than the time between successive immigration events, however. When $N\nu \gg 1$ the system is still evolving when a new immigrant is introduced, which acts to keep the probability distribution near g and away

from fixation. In this limit the long term variance tends to $\sigma^2(\infty) \approx g(1-g)N/\nu$. The argument for having no variance with $g = 0, 1$ still stands. But now that the immigration rate is no longer negligibly small, it shows up in the variance. For a fixed system size N , increasing the immigration rate decreases the variance, as the system is drawn more toward the metapopulation abundance and away from the edges.

The variance limits, and indeed figure 1.4, suggest that there are two regimes of the Moran model with immigration. At low immigration rate the system undergoes a series of fixations punctuated by the occasional immigrant. It spends most of its time resting in the fixated state, rarely seeing a new immigrant which quickly either dies out or takes over in a new fixation. When immigration is common the system is tied to the metapopulation, and deviations away from the metapopulation abundance are suppressed. Probability gathers near the mean value gN . These regimes will be investigated further in the following paragraphs.

A quantity similar to the mean is the extremum of the distribution, which for large immigration corresponds to the mode of the system. The extremum occurs when $\partial_n \tilde{P}_n = 0$ but for ease note that $\partial_n \tilde{P}_n = \partial_n q_n / \sum_i q_i = \partial_n q_n = q_n \partial_n \ln(q_n)$ therefore I can instead calculate the value which gives $\partial_n \ln(q_n) = 0$. First,

$$\ln(q_n) = 2 \ln[N] - \ln[n(N-n)(1-\nu) + (1-g)nN\nu] + \ln[(N-n)!]$$

Also finding the mode and inversion values...!!!

Figure 1.4 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

$$E_i = \frac{\sum_{j=i}^{N-1} S_j}{1 + \sum_{j=1}^{N-1} S_j}. \quad (1.13)$$

See figure 1.5 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. Comment here? Or later...!!!

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

$$\tau[i] = \sum_{k=1}^{N-1} q_k + \sum_{j=1}^{i-1} S_j \sum_{k=j+1}^{N-1} q_k. \quad (1.14)$$

Note that this should go to zero at both $n = 0$ and $n = N$, since it is unconditioned. Again, there is a closed form, but it is a sum of hyperbolic functions and does not possess intuitive limits. It is approximated numerically and displayed graphically in figure 1.6. Comments???!!!

Keeping with the artificial stoppage when the focal population reaches 0 or N individuals, we calculate

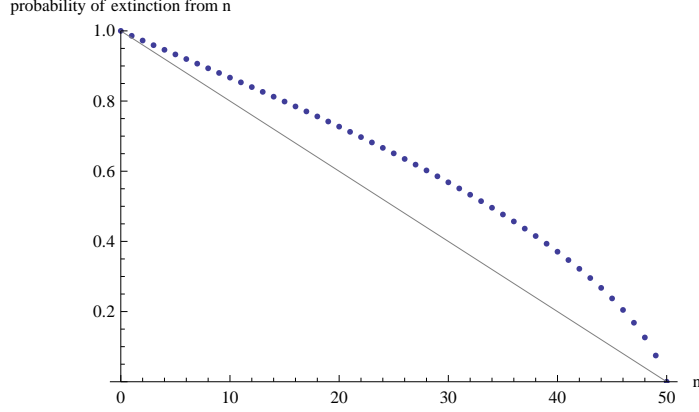


Figure 1.5: Probability of first going extinct, given starting population/fraction. $g = 0.1$, $N = 50$, $\nu = 0.01$. Grey is regular Moran results without immigration.

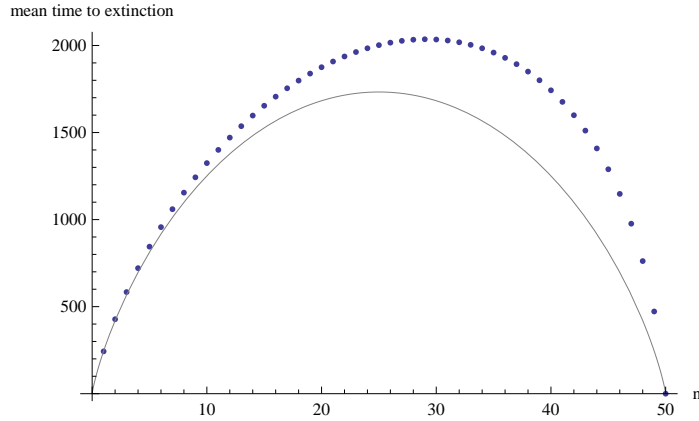


Figure 1.6: Mean time to either fixation or extinction, given starting population/fraction. $g = 0.1$, $N = 50$, $\nu = 0.01$. Grey is regular Moran results without immigration.

the conditional times, respectively to extinction and to fixation. As before, the extinction probability is given by equation 1.13. Similar to the continuous time solutions presented in the introduction, the conditional extinction time can be written as

$$\phi[n] = \phi[1] + \sum_{j=1}^{n-1} \left(\phi[1] - \sum_{i=1}^j q_i E_i \right) S_j. \quad (1.15)$$

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

$$\phi_1 = \frac{\sum_{j=1}^{N-1} \sum_{i=1}^j q_i E_i}{1 + \sum_{j=1}^{N-1} S_j}. \quad (1.16)$$

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 1.7. The conditional times mostly follow the unconditioned time, except near the rare events that do not much contribute to the average.

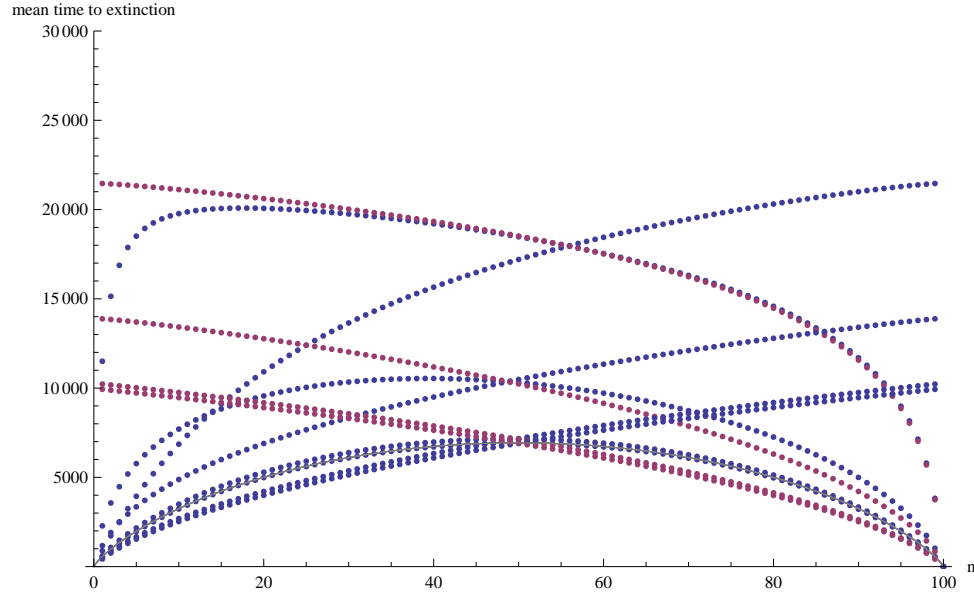


Figure 1.7: Mean time to fixation or extinction, conditioned on that event happening, given starting population/fraction. $g = 0.7$, $N = 100$, ν varies from 0.3 (highest) to 0.0001 (lowest). Grey is regular Moran results without immigration.

1.6 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.