I. Moran model

The Moran model describes the evolution of a collection of cells that is maintained at a constant population size of N. Thinking of the population genetics situation for which it was developed, we will often call the cells individuals. In comparing the results here with those given in the literature one must take into account that in genetics most organzisms are diploid (have two copies of their DNA), so the population in the Moran model commonly consists of $2N$ copies of a locus.

1 Neutral case

In the simplest version of the model (with no selection or mutation), the dynamics of the Moran model, which occur in continuous time, can be described as follows:

- Each individual is replaced at rate 1. That is, individual x lives for an exponentially distributed amount with mean 1 and then is "replaced."
- To replace individual x, we choose an individual at random from the population (including x itself) to be the parent of the new individual.

Suppose now that each individual has one of two alleles A and a, and let X_t be the number of copies of A. The transition rates for X_t are

$$\begin{aligned} i \to i+1 \quad &\text{at rate} \quad b_i = (N-i)\cdot \frac{i}{N} \\ i \to i-1 \quad &\text{at rate} \quad d_i = i \cdot \frac{N-i}{N} \end{aligned} \tag{1}$$

where b is for birth and d is for death. In words, a's are selected for possible replacement at total rate $N-i$. The number of A's will increase if an A is chosen to be the parent of the new individual, an event of probability i/N. Similarly, A's are selected for possible replacement at total rate i. The number of A's will decrease if an a is chosen to be the parent of the new individual, an event of probability $(N-i)/N$. Note that $b_i = d_i$.

Let $\tau = \min\{t : X_t = 0 \text{ or } X_t = N\}$ be the fixation time, i.e., the first time at which all individuals have the same type. Since it is possible to reach the absorbing states 0 and N starting from any interior state $0 < i < N$ we have $P_i(\tau < \infty) = 1$

Theorem 1. *In the Moran model, the probability that A becomes fixed when there are initially i copies is i/N.*

Proof. The rates for up and down jumps are the same, so $(d/dt)E_iX_t = 0$, and hence E_iX_t is constant, i.e., X_t is a martingale. Intuitively this implies that

$$i = E_iX_\tau = NP_i(X_\tau = N) \tag{2}$$

To prove this we note that

$$i = E_iX_t = NP_i(X_\tau = N, \tau \le t) + E_i(X_t; \tau > t)$$

Letting $t \to \infty$ and noting $P_i(\tau > t) \to 0$, $|X_t| \le N$ the desired result follows. □

Let $T_k = \min\{t : X_t = k\}$ be the hitting time of k. Writing $\bar{E}_i\tau = E_i(\tau|T_N < T_0)$, we can state

Theorem 2. *Let $p = i/N$. In the Moran model when N is large*

$$\bar{E}_i\tau \approx -\frac{N(1-p)}{p}\log(1-p) \tag{3}$$

As $p \to 0$, $-\log(1-p)/p \to 1$, so

$$\bar{E}_1\tau \approx 2N \tag{4}$$

Proof. Let S_j be the amount of time spent at j before time τ and note that

$$E_i\tau = \sum_{j=1}^{N-1} E_iS_j \tag{5}$$

Let N_j be the number of visits to j. Let $q(j) = 2j(N-j)/N$ be the rate at which the chain leaves j. Since each visit to j lasts for an exponential amount of time with mean $1/q(j)$, we have

$$E_iS_j = \frac{1}{q(j)}E_iN_j \tag{6}$$

To compute E_iN_j, we begin by noting that

$$P_i(N_j \geq 1) = P_i(T_j < \infty)$$

Letting $T_j^+ = \min\{t : X_t = j \text{ and } X_s \neq j \text{ for some } s < t\}$ be the time of the first return to j, we have for $n \geq 1$

$$P_i(N_j \geq n+1 | N_j \geq n) = P_j(T_j^+ < \infty)$$

The last formula shows that, conditional on $N_j \geq 1$, N_j has a geometric distribution with success probability $P_j(T_j^+ = \infty)$. Combining this with our formula for $P_i(N_j \geq 1)$, we have

$$E_iN_j = \frac{P_i(T_j < \infty)}{P_j(T_j^+ = \infty)} \tag{7}$$

Since the average value of X_t is constant in time, the martingale argument in (2) shows that for $0 \leq i \leq j$

$$i = jP_i(T_j < T_0) + 0 \cdot [1 - P_i(T_j < T_0)]$$

and solving gives

$$P_i(T_j < T_0) = \frac{i}{j} \qquad P_i(T_j > T_0) = \frac{j-i}{j} \tag{8}$$

Similar reasoning shows that for $j \leq i \leq N$,

$$i = jP_i(T_j < T_N) + N[1 - P_i(T_j < T_N)]$$

and solving gives

$$P_i(T_j < T_N) = \frac{N-i}{N-j} \qquad P_i(T_j > T_N) = \frac{i-j}{N-j} \tag{9}$$

When the process leaves j, it goes to $j-1$ or $j+1$ with equal probability, so

$$\begin{aligned} P_j(T_j^+ = \infty) &= \frac{1}{2} \cdot P_{j+1}(T_j > T_N) + \frac{1}{2} \cdot P_{j-1}(T_j > T_0) \\ &= \frac{1}{2} \cdot \frac{1}{N-j} + \frac{1}{2} \cdot \frac{1}{j} = \frac{N}{2j(N-j)} \end{aligned}$$

Putting our results into (7) gives

$$E_i N_j = \begin{cases} \frac{i}{j} \cdot \frac{2j(N-j)}{N} & i \le j \\ \frac{N-i}{N-j} \cdot \frac{2j(N-j)}{N} & j \le i \end{cases}$$

Since $q(j) = 2j(N-j)/N$, (6) gives

$$E_i S_j = \begin{cases} \frac{i}{j} & i \le j \\ \frac{N-i}{N-j} & j \le i \end{cases} \tag{10}$$

If we let $h(i) = P_i(T_N < T_0)$ and let $p_t(i,j)$ be the transition probability for the Moran model, then it follows from the definition of conditional probability and the Markov property that

$$\bar{p}_t(i,j) = \frac{P_i(X_t = j, T_N < T_0)}{P_i(T_N < T_0)} = p_t(i,j) \cdot \frac{h(j)}{h(i)}$$

Integrating from $t = 0$ to ∞, we see that the conditioned chain has

$$\bar{E}_i S_j = \int_0^\infty \bar{p}_t(i,j)\, dt = \frac{h(j)}{h(i)} E_i S_j \tag{11}$$

$h(i) = i/N$, so $h(j)/h(i) = j/i$ and using the formula for $E_i S_j$ given in (10), we have

$$\bar{E}_i S_j = \begin{cases} 1 & i \le j \\ \frac{j}{i} \cdot \frac{N-i}{N-j} & j \le i \end{cases} \tag{12}$$

By the reasoning that led to (5),

$$\bar{E}_i \tau = \sum_{j=1}^{N-1} \bar{E}_i S_j = \sum_{j=i}^{N-1} 1 + \frac{N-i}{i} \cdot \sum_{j=1}^{i-1} \frac{j}{N-j}$$

The first sum is $N-i$. For the second we note that

$$\sum_{j=1}^{i-1} \frac{j}{N-j} = N \sum_{j=1}^{i-1} \frac{j/N}{1-j/N} \cdot \frac{1}{N} \approx N \int_0^p \frac{u}{1-u}\, du$$

where $p = i/N$. To evaluate the integral, we note that it is

$$= \int_0^p -1 + \frac{1}{1-u}\, du = -p - \log(1-p)$$

Combining the last three formulas gives

$$\begin{aligned}\bar{E}_i\tau &\approx N(1-p) + \frac{N(1-p)}{p}(-p - \log(1-p)) \\ &= -\frac{N(1-p)}{p}\log(1-p)\end{aligned}$$

which gives (3). □

2 Directional selection

In this section, we will introduce selection letting 1 and $1-s$ be the relative fitnesses of the two alleles, A and a. Let X_t be the number of A's at time t. Thinking of the fitnesses as the probability that an offspring of that type is viable, we can formulate the transition rates of the Moran model with selection as

$$\begin{aligned} i &\to i+1 \quad \text{at rate } b_i = (N-i)\cdot\frac{i}{N} \\ i &\to i-1 \quad \text{at rate } d_i = i\cdot\frac{N-i}{N}\cdot(1-s) \end{aligned} \tag{13}$$

In words, a's are selected for possible replacement at total rate $N-i$. The number of A's will increase if an A is chosen to be the parent of the new individual, an event of probability i/N. The reasoning is similar for the second rate, but in this case the replacement only occurs with probability $1-s$.

Theorem 3. *In the Moran model with selection $s > 0$*

$$P_i(T_N < T_0) = \frac{1-(1-s)^i}{1-(1-s)^N} \tag{14}$$

When $i = 1$, the numerator is just s. If selection is strong, i.e., Ns is large, then $(1-s)^N \approx 0$ and the probability of fixation of a new mutant is just s. When $Ns = O(1)$, $(1-s) \approx e^{-s}$, so (14) can be written as

$$P_i(T_N < T_0) \approx \frac{1-e^{-is}}{1-e^{-Ns}} \tag{15}$$

Proof. Let $h(i) = P - i(T_N < T_0)$. Births happen at rate b_i and deaths at rate d_i, so the probability a birth occurs before a death is $b_i/(b_i+d_i)$ and we have

$$h(i) = \frac{b_i}{b_i+d_i}h(i+1) + \frac{d_i}{b_i+d_i}h(i-1)$$

Multiplying on each side by $b_i + d_i$ and rearranging, we have

$$h(i+1) - h(i) = \frac{d_i}{b_i}(h(i) - h(i-1)) = (1-s)(h(i) - h(i-1))$$

Now $h(0) = 0$, so if we let $c = h(1)$ and iterate, it follows that

$$(\star) \qquad h(i+1) - h(i) = c(1-s)^i$$

Summing we have

$$h(j) = \sum_{i=0}^{j-1} c(1-s)^i = c\frac{1-(1-s)^j}{s}$$

We must have $h(N) = 1$ so $c = s/(1-(1-s)^N)$ and the desired result follows. □

We can also prove the result using a more intuitive martingale argument.

Another derivation of (14). To motivate the computation, we begin by recalling the martingale proof of Theorem 1. Let $\tau = T_0 \wedge T_N$. When $s = 0$, EX_t is constant in time, so we have

$$i = N \cdot P_i(X_\tau = N) + 0 \cdot P_i(X_\tau = 0)$$

Solving, we have $P_i(X_\tau = N) = i/N$. When $s > 0$, $b_i/(b_i + d_i) = 1/(2-s)$. A little calculation shows that

$$\begin{aligned}(1-s)^{i+1}\frac{1}{2-s} + (1-s)^{i-1}\frac{1-s}{2-s} \\ = (1-s)^i\frac{1-s}{2-s} + (1-s)^i\frac{1}{2-s} = (1-s)^i\end{aligned}$$

so, in this case, the value of $E(1-s)^{X_t}$ stays constant in time. Reasoning as before,

$$(1-s)^i = (1-s)^N P_i(X_\tau = N) + 1 \cdot [1 - P_i(X_\tau = N)]$$

Solving we have

$$P_i(X_\tau = N) = \frac{1-(1-s)^i}{1-(1-s)^N}$$

in agreement with (14). □

One can generalize Theorem 2 to compute the expected time to fixation in the model with selection. However to obtain more insight into what is happening during the fixation of a favorable allele, we will take a different approach.

Theorem 4. *In the Moran model with selection as $N \to \infty$*

$$\bar{E}_1\tau \sim \frac{2}{s}\log N$$

Proof. The key is to establish that there are three phases in the fixation process.

1. While the advantageous A allele is rare, the number of A's can be approximated by a supercritical branching process.

2. While the frequency of A's, $u_A \in [\epsilon, 1-\epsilon]$ there is very little randomness and u_A follows the solution of the logistic differential equation: $du_A/dt = su_A(1-u_A)$.

3. While the disadvantageous a allele is rare, the number of a's can be approximated by a subcritical branching process.

Phase 1. Let i be the number of A's. If i/N is small, then the transition rates in (13) simplify:

$$\begin{aligned} i &\to i+1 \quad \text{at rate } b_i \approx i \\ i &\to i-1 \quad \text{at rate } d_i \approx (1-s)i \end{aligned}$$

This is a continuous time branching process in which each of the i individuals gives birth at rate 1 and dies at rate $1-s$. Letting Z_t be the number of individuals at time t, it is easy to see from the description that

$$\frac{d}{dt} EZ_t = sEZ_t$$

so $EZ_t = Z_0 e^{st}$. A result from the theory of branching processes, see Athreya and Ney (1972), shows that as $t \to \infty$

$$e^{-st} Z_t \to W \tag{16}$$

The limit W may be 0, and will be if the branching process dies out, that is, if $Z_t = 0$ for some t. However, on the event that the process does not die out $\Omega_\infty = \{Z_t > 0 \text{ for all } t\}$, we have $W > 0$.

Let T_1 be the first time that $X_t = M = N/\log N$. Using (16), we see that $(e^{-st} Z_t | \Omega_\infty) \to \bar{W} = (W | W > 0)$ so if we condition on survival

$$\frac{N}{\log N} \approx \exp(sT_1)\bar{W}$$

and solving gives

$$T_1 \approx \frac{1}{s} \log\left(\frac{N}{\bar{W} \log N}\right) \approx \frac{1}{s} \log(N)$$

Phase 2. Let T_2 be the first time that $X_t = N - M$, where $M = N/\log N$. As we will now show, during the second phase from T_1 to T_2 the process behaves like the solution of the logistic differential equation. Let X_t be the number of copies of the mutant allele at time t, and let $Y_t^N = X_t/N$. Y_t^N makes transitions as follows:

$$\begin{aligned} i/N &\to (i+1)/N \quad \text{at rate } b_i = N - i \cdot \frac{i}{N} \\ i/N &\to (i-1)/N \quad \text{at rate } d_i \approx (1-s)i \cdot \frac{N-i}{N} \end{aligned}$$

When $Y_0^N = i/N = y$, the infinitesimal mean

$$\frac{d}{dt} EY_t^N = b_i \cdot \frac{1}{N} + d_i \cdot \left(-\frac{1}{N}\right) = s\frac{N-i}{N} \cdot \frac{i}{N} = sy(1-y)$$

while the infinitesimal variance

$$\frac{d}{dt} E(Y_t^N - y_0)^2 = (b_i + d_i) \cdot \frac{1}{N^2} = (2-s)\frac{N-i}{N} \cdot \frac{i}{N} \cdot \frac{1}{N} \to 0$$

In this situation, results in Section 7.4 of Ethier and Kurtz (1986), show that as $N \to \infty$, Y_t^N converges to Y_t, the solution of the logistic differential equation

$$dY_t = sY_t(1 - Y_t)$$

It is straightforward to check that the solution of this equation is

$$Y_t = \frac{1}{1 + Ce^{-t}}$$

where $C = (1 - Y_0)/Y_0$. In the case of interest, $Y_0 = 1/\log(N)$, so $C \approx \log(N)$.Thus $Y_t = 1 - 1/(\log N)$ when

$$(\log N)e^{-t} = \frac{\log N}{\log N - 1} - 1 = \frac{1}{\log N - 1} \sim \frac{1}{\log N}$$

Solving, we find that $T_2 - T_1 \approx 2 \log\log N$.

Phase 3. To achieve fixation of the A allele mutation after time T_2, the $M = N/(\log N)$ a alleles must decrease to 0. The number of a alleles, Z_t, makes transitions

$$\begin{aligned} j &\to j + 1 \quad \text{at rate } d_{N-j} \approx (1 - s)j \\ j &\to j - 1 \quad \text{at rate } b_{N-j} \approx j \end{aligned}$$

That is, Z_t is a continuous time branching process in which each of the j individuals gives birth at rate $(1 - s)$ and dies at rate 1. By arguments in phase 1, $EZ_t = Z_0 e^{-st}$ so it takes about $(1/s)\log(2N)$ units of time to reach 0.

The times in the three phases were

Phase 1	$(1/s)\log(N)$
Phase 2	$\log\log(N)$
Phase 3	$(1/s)\log(N)$

and we have proved Theorem 4. □

3 Waiting for two mutations

Consider now a version of the Moran model in which initially all cells are type 0, and where in addition to the usual replacement dynamics, cells of type $i-1$ mutate to type i at rate u_i. In some treatments mutations are only allowed to occur at births, but this small detail makes very little difference to results, so we will take the mathematically simpler approach of having mutation as a separate process.

We are interested in τ_k the time the first type-k individual occurs. There are at least three reasons to be interested in the case $k = 2$.

Tumor suppressor genes, for example *APC* in colon cancer or *RB*1 in retinoblastoma directly regulate cellualr growth and differentiation. Inactivation of one copy does not lead to a phenotypic change, but loss of both can increase the net reproductive rate of the call and represent a step toward cancer.

Regulatory sequence evolution. There is a growing body of experimental evidence that in Drosophila, significant changes in gene regulation can occur in a short amount of time, compared to divergence time between species. Ludwig et al. (1998, 2000, 2005) studied the evolution of the *even-skipped* stripe 2 enhancer in four Drosophila species (*D. melanogaster*, *D. yakuba*, *D. erecta*, and *D. pseuodoobscura*). While expression is strongly conserved, they found many substitutions in the binding sites for bicoid, hunchback, Kruppel, and giant, as well as large differences in the overall size of the enhancer region. In this context the first mutation might damage an exisitng binding site while the second one mutation creates a new binding site or the events could occur in the other order. However, we can estimate the rate of this type of evolution if we can analyze the situation when we are looking for two mutation in a prescribed order.

The edge of evolution? Our final example of waiting for two mutations concerns the emergence of chloroquine resistance in *P. falciparum*. Genetic studies have shown, see Wooton *et al.* (2002), that this is due to changes in a protein PfCRT, and that in the mutant strains two amino acid changes are almost always present - one switch at position 76 and another at position 220. This example plays a key role in the chapter titled "The Mathematical Limits of Darwinism" in Michael Behe's book, *The Edge of Evolution*.

Arguing that (i) there are a trillion parasitic cells in an infected person, (ii) a billion infected persons on the planet, and (ii) chloroquine resistance has only arisen ten times in the last fifty years, he concludes that the odds of one parasite developing resistance to chloroquine, an event he calls a *chloroquine complexity cluster* or CCC, is roughly 1 in 10^{20}. Ignoring the fact that humans and *P. falciparum* have different mutation rates he then concludes that "On the average, for humans to achieve a mutation like this by chance, we would have to wait a hundred million times ten million years."

To further sensationalize his conclusion, he argues that "There are 5000 species of modern mammals. If each species had an average of a million members, and if a new generation appeared each year, and if this went on for two hundred million years, the likelihood of a single CCC appearing in the whole bunch over that entire time would only be about 1 in 100." In our contrast our result predicts a waiting time of 31.6 million generations for one prespecified pair of mutations in one species.

3.1 Stochastic tunneling

One boring scenario for producing a type 2 is that a type 1 mutation occurs and fixes in the population, then a mutation to type 2 occurs. Writing $a \ll b$ as short for a/b is small or more precisely the assumption that $a_N/b_N \to 0$, there is the following more interesting possibility

Theorem 5. *If* $1/\sqrt{u_2} \ll N \ll 1/u_1$ *then as* $N \to \infty$

$$P(\tau_2 > t/Nu_1\sqrt{u_2}) \to e^{-t}$$

To begin to explain the intuition that underlies this result, let r be the probability that a type 1 gives birth to a type 2 before its family line dies out. By considering what happens at the first event we see that

$$r = \frac{u_2}{u_2+2} + \frac{1}{u_2+2}(2r - r^2) \tag{17}$$

To check this, it is convenient to write A for type-1 and B for type-2 then we have

$$\begin{aligned} A &\to 0 \quad \text{at rate } \frac{1\cdot(N-1)}{N} \approx 1 \\ A &\to 2A \quad \text{at rate } \frac{(N-1)\cdot 1}{N} \approx 1 \\ A &\to B \quad \text{at rate } u_2 \end{aligned}$$

The mutation $A \to B$ will occur first with probability $u_2/(u_2+2)$ in which case success is assured (i.e., we will get a 2). Removal of the lone A, $A \to 0$, will occur first with probability $1/(u_2+2)$ in which case success is impossible. Finally, $A \to 2A$ will occur first with probability $1/(u_2+2)$ in which case success has probability $1-(1-r)^2 = 2r - r^2$ since successes for the the two lineages is almost independent.

A little algebra converts (17) into $r^2 + u_2 r - u_2 = 0$ which has positive solution

$$\frac{-u_2 + \sqrt{u_2^2 + 4u_2}}{2} \approx \sqrt{u_2} \tag{18}$$

since $u_2^2 \ll u_2 \ll \sqrt{u_2}$ when u_2 is small. While the number of 1's in the population remains $o(N)$, mutations to type 1 occur at rate Nu_1. By (18) the probability that a type 1 mutation will give rise to a type 2 is $\sim \sqrt{u_2}$. Using Poisson thinning now we see that if σ_2 is the time of the first type 1 mutation that gives rise to a type 2 then σ_2 is approximately exponential with rate $Nu_1\sqrt{u_2}$.

Up to this point we have not used the assumptions of the theorem. To explain how they enter the picture, we note

1. The number of type 1's is a time change of a symmetric random walk, so if the number reaches M then there will be of order M^2 births before the family of type 1's dies out. From this we see that the type 2 mutation will first occur in a family that reaches size $O(1/\sqrt{u_2})$. If we want our assumption about the number of type 1's up to τ_2 to be $o(N)$ then we must have $1/\sqrt{u_2} \ll N$.

2. We have a limit theorem for σ_2 but we want one for τ_2, so we need to show that $\tau_2 - \sigma_2$ can be neglected. To do this we note that Theorem 2 implies

$$E_1(T_M | T_M < T_0) \sim M$$

We have $M = O(1/\sqrt{u_2})$ while $E\sigma_2 = O(1/nu_1\sqrt{u_2})$ so for $\tau_2 - \sigma_2 = o(E\sigma_2)$ we need $Nu_1 \ll 1$.

3.2 Proof of Theorem 5

Our next goal is to prove a result that is a little more general than Theorem 5 in that it allows $Nu_1 \to \lambda > 0$.

Theorem 6. *Suppose that $Nu_1 \to \lambda \in [0, \infty)$, $u_2 \to 0$, and $N\sqrt{u_2} \to \infty$ as $N \to \infty$. Then*

$$P\left(\tau_2 > t/Nu_1\sqrt{u_2}\right) \to \exp\left(-\int_0^t h(s)\, ds\right)$$

where $h(s) = (1 - e^{-2s/\lambda})/(1 + e^{-2s/\lambda})$ if $\lambda > 0$ and $h(s) \equiv 1$ if $\lambda = 0$.

If we let $X_1(t)$ be the number of type 1 individuals at time t then

$$P(\tau_2 > t) = E \exp\left(-u_2 \int_0^t X_1(s)\, ds\right) \tag{19}$$

Step 1. *We can replace $X_1(t)$ by a continuous-time critical branching process, $Y(t)$, with births and deaths at rate 1 and immigration at rate Nu_1.*

When $X_1(t) = k$, type 1 mutations occur at rate $(N-k)u_1$, while birth events in which a type 1 individual replaces a type 0 individual occur at rate $k(N-k)/N$, so we have jumps

$$\begin{aligned} k &\to k+1 \quad \text{at rate } (k + Nu_1) \cdot \frac{N-k}{N} \\ k &\to k-1 \quad \text{at rate } k \cdot \frac{N-k}{N} \end{aligned}$$

The branching process with immigration, $Y(t)$, has jumps

$$\begin{aligned} k &\to k+1 \quad \text{at rate } k + Nu_1 \\ k &\to k-1 \quad \text{at rate } k \end{aligned}$$

Comparing rates we see that the process $\{X_1(t), t \ge 0\}$ is a time-change of $\{Y(t), t \ge 0\}$, in which time runs slower than in the branching process by a factor of $(N-k)/N$. That is if

$$T(t) = \int_0^t \frac{N - X_1(s)}{N}\, ds \le t$$

then the two processes can be coupled so that $X_1(t) = Y(T(t))$, for all $t \ge 0$. The time change will have little effect as long as $X_1(t)$ is $o(N)$.

Step 2. *On the relevant time scale, the number of 1s stays small with high probability.*

Lemma 1. *Fix $t > 0$, $\epsilon > 0$, and let $M_t = \max_{0 \le s \le t/(Nu_1\sqrt{u_2})} X_1(s)$. We have*

$$\lim_{N\to\infty} P\left(M_t > \epsilon N \right) = 0.$$

Proof. Since in addition to the immigration, individuals give birth and die at the same rate, the process $\{X_1(s), s \ge 0\}$ is a submartingale. Because the rate of type 1 mutations is always bounded above by Nu_1, we have $EX_1(s) \le Nu_1 s$ for all s. By Doob's Maximal Inequality,

$$P(M_t > \epsilon N) \le \frac{EX_1(t/Nu_1\sqrt{u_2})}{\epsilon N} \le \frac{1}{\epsilon N} \cdot \frac{Nu_1 t}{Nu_1\sqrt{u_2}},$$

which goes to zero as $N \to \infty$, since $N\sqrt{u_2} \to \infty$. □

Step 3. *A useful lemma.*

Steps 1 and 2 have shown that it is enough to prove the result for the branching process, $Y(t)$. Let Q denote the distribution of $\{Y(t), t \ge 0\}$, and let Q_1 denote the law of the process starting from a single type 1 and modified to have no further mutations to type 1.

Lemma 2. *The waiting time for the first type 2 in a system with type 1 mutations at rate Nu_1 satisfies*

$$Q(\tau_2 \le t) = 1 - \exp\left(-Nu_1 \int_0^t Q_1(\tau_2 \le s)\, ds \right) \tag{20}$$

Proof. Type 1 mutations are a Poisson process with rate Nu_1. A point at time $t - s$ is a success, i.e., produces a type 2 before time t with probability $Q_1(\tau_2 \le s)$. By results for thinning a Poisson process, the number of successes by time t is Poisson with mean $Nu_1 \int_0^t Q_1(\tau_2 \le s)\, ds$. The result follows from the observation that $Q(\tau_2 \le t)$ is the probability of at least one success in the Poisson process. □

Step 4. *Compute $g_2(t) = Q_1(\tau_2 \le t)$.*

By considering what happens between time 0 and h

$$g_2(t+h) = g_2(t)[1 - (2+u_2)h] + h[2g_2(t) - g_2(t)^2] + h \cdot 0 + u_2 h \cdot 1 + o(h)$$

where the four terms correspond to nothing happening, a birth, a death, and a mutation of the original type 1 to type 2. Doing some algebra and letting $h \to 0$

$$g_2'(t) = -u_2 g_2(t) - g_2(t)^2 + u_2 \tag{21}$$

If we let $r_1 > r_2$ be the solutions of $x^2 + u_2 x - u_2 = 0$, i.e.,

$$r_i = \frac{-u_2 \pm \sqrt{u_2^2 + 4u_2}}{2} \tag{22}$$

we can write this as $g_2'(t) = -(g_2(t) - r_1)(g_2(t) - r_2)$. A little calculus gives

$$g_2(t) = \frac{r_1(1 - e^{(r_2 - r_1)t})}{1 - (r_1/r_2)e^{(r_2 - r_1)t}}$$

Using the facts that $r_1 - r_2 = \sqrt{u_2^2 + 4u_2} \sim 2\sqrt{u_2}$ and $r_1/r_2 \to -1$ we see that if $t\sqrt{u_2} \to s$ then

$$g_2(t) \sim \sqrt{u_2} \cdot \frac{1 - e^{-2s}}{1 + e^{-2s}} \tag{23}$$

Using (23) in Theorem 2 gives the result in Theorem 6.

3.3 The assumption $N\sqrt{u_2} \to \infty$

In the previous subsection we saw that the assumption $Nu_1 \to 0$ was needed to be able to ignore the difference $\tau_2 - \sigma_2$ in the limit theorem. The assumption $N\sqrt{u_2} \to \infty$ implies that the 1's are $o(N)$ until time τ_2. In the other direction,

Theorem 7. *If $Nu_1 \to 0$ and $N\sqrt{u_2} \to 0$ then $\tau_2 \Rightarrow$ exponential(u_1) + exponential(Nu_2), the sum of two independent exponentials.*

Proof. If $N\sqrt{u_2} \ll 1$ then the probability of fixation $1/N \gg \sqrt{u_2}$ the probability a type 1 mutant gives rise to a type 2 before its family does out. From this we see that with high probability a type 1 mutation destined for fixation will occur before the first type 2 mutation occurs. Type 1 mutations that fix occur at rate $Nu_1 \cdot 1/N$. By Theorem 2 the average waiting time for fixation conditioned that it occurs is $\sim N \ll 1/u_1$ so this can be ignored. One the 1's fix the waiting time for the mutation to type 2 is exponential(Nu_2). □

There is interesting behavior in borderline case between Theorem 7 and 5.

Theorem 8. *Suppose that $Nu_1 \to 0$ and $(Nu_1\sqrt{u_2})^2 \to \gamma > 0$, and let*

$$\alpha = \sum_{k=1}^{\infty} \frac{\gamma^k}{(k-1)!(k-1)!} \Big/ \sum_{k=1}^{\infty} \frac{\gamma^k}{k!(k-1)!} > 1 \tag{24}$$

then $P(u_1\tau_2 > t) \to \exp(-\alpha t)$.

In this case the mutation to type 2 occurs with positive probability before fixation occurs, and at a time when then number of type 1's, $X_1(t)$, is $O(N)$. The keys to the analysis are

- If we start with $X_1(0) = N\epsilon$ then $N^{-1}X_1(Nt)$ converges to Y_t where Y_t is the Wright-Fisher diffusion with infinitesimal generator $x(1-x)d^2/dx^2$.
- If $Y(0) = y$ and $u(y)$ is the probability that starting from $X_1(Nt) = Ny$ the process hits 0 before reaching 1 or generating a type 2 mutation satisfies

$$x(1-x)u''(x) - \gamma x u(x) \qquad u(0) = u(1) = 1 \tag{25}$$

The solution of (25) is

$$u(y) = c\sum_{k=1}^{\infty} \frac{\gamma^k}{k!(k-1)!}(1-y)^k$$

and the constant $\alpha = u'(0)$. The details are somwhat complicated tso we refer the reader to Durrett, Schmidt, and Schweinsberg.

References

Behe, M., 2007 *The Edge of Evolution. The search for the limits of Darwinism.* Free Press, New York.

Durrett, R., Schmidt, D., and Schweinsberg, J. (2007) A waiting time problem arising from the study of multi-stage carcinogenesis. *Annals of Applied Probability*

Iwasa, Y., Michor, F., and Nowak, M.A. (2004) Stochastic tunnels in evolutionary dynamics. Genetics **166:** 1571–1579.

Iwasa, Y., Michor, F., Komarova, N.L., and Nowak, M.A. (2005) Population genetics of tumor suppressor genes. J. Theor. Biol. 233. 15–23

Ludwig, M. Z., Patel, N.H., and Kreitman, M. (1998) Functional analysis of eve stripe 2 enhancer evolution in Drosophila: rules governing conservation and change. *Development.* 125, 949–958

Ludwig, M. Z., Bergman, C.E., Patel, N.H., and Kreitman, M. (2000) Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* 403, 564–567

Ludwig, M. Z., Palsson, A., Alekseeva, E., Bergman, C.E., Nathan, J., and Kreitman, M. (2005) Functional evolution of a *cis*-regulatory module. *PLoS Biology.* 3, 588–598

Nowak, M. A. (2006) *Evolutionary Dynamics: Exploring the Equations of Life.* Belknap Press, Cambridge, MA

Nowak, M.A., Michor, F., Komorova, N.L., and Iwasa, Y. (2004) Evolutionary dyanmics of tumor suppressor gene inactivation. *Proc. Natl.Acad. Sci.* 101, 10635–10638

Wooton, J. C., Feng, A., Ferdig, M. T., Cooper, R. A., Mu, J., et al., (2002) Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum. Nature.* 418, 320–323

Wodarz, D., and Komarova, N.L. (2005) *Computational Biology of Cancer: Lecture notes and mathematical modeling.* World Scientific Publishing, Hackensack, NJ.