0.1 General Model

Let the type of an individual be the number of mutations and let $p_{ki}(j) = P_k(X^{(i)} = j)$, the probability that a k-type gets j i-type offspring for $i = k, k+1, ...$

Let $p_k(j)$ be the probability that a k-type individual gets j offspring (so $p_k(j) = \sum_i p_{ki}(j)$).

Let q_r be the probability that an offspring gets r new mutations. Then

$$p_{ki}(j) = \sum_{l=j}^{\infty} \binom{l}{j} q_{i-k}^j (1 - q_{i-k})^{l-j} p_k(l)$$
$$= (\frac{q_{i-k}}{1 - q_{i-k}})^j \sum_{l=j}^{\infty} \binom{l}{j} (1 - q_{i-k})^l p_k(l).$$

This has pgf

$$\varphi_{ki}(t) = \sum_{j=0}^{\infty} t^j p_{ki}(j) = \sum_{j=0}^{\infty} t^j \sum_{l=j}^{\infty} \binom{l}{j} q_{i-k}^j (1 - q_{i-k})^{l-j} p_k(l)$$
$$= \sum_{l=0}^{\infty} (1 - q_{i-k} + tq_{i-k})^l p_k(l) = \varphi_k(1 - q_{i-k} + tq_{i-k})$$

where $\varphi_k(t) = \sum_{j=0}^{\infty} t^j p_k(j)$, the pgf of the number of offspring of a k-type.

Assuming reproducion is binomial(N, p), we get

$$\varphi_{ki}(t) = (1 - p + p(1 - q_{i-k} + tq_{i-k}))^N$$

and

$$m_{ik} = Npq_{i-k}.$$

0.2 Our model

An individual of type k splits with probability $pw(k)$, stays with probability $qw(k)$, and dies with probability $1 - (p + q)w(k)$, where the fitness function

$w(k)$ has $w(0) = 1$. Any offspring may acquire one mutation with probability λ, independently. Thus, the number of mutant offspring of a surviving individual has a binomial distribution with parameters 2 and λ. The possible reproduction schemes are:

$$k \to \begin{cases} k \\ k \quad \text{with probability } pw(k)(1-\lambda)^2 \end{cases}$$

$$k \to \begin{cases} k \\ k+1 \quad \text{with probability } 2pw(k)\lambda(1-\lambda) \end{cases}$$

$$k \to \begin{cases} k+1 \\ k+1 \quad \text{with probability } pw(k)\lambda^2 \end{cases}$$

$$k \to k \text{ with probability } qw(k)$$

$$k \to \emptyset \text{ with probability } 1-(p+q)w(k)$$

The mean reproduction matrix M has entries

$$\begin{cases} m_{k,k} &= (2p(1-\lambda)+q)w(k) \\ \\ m_{k,k+1} &= 2p\lambda w(k) \end{cases}$$

for $k = 0, 1, 2, \ldots$, all other entries being 0. The mean reproduction matrix for the nth generation is M^n, the nth power of M. Note that, for any n, all entries of M^n below the diagonal are 0. If fitness is multiplicative, $w(k) = (1-s)^k$ and we can get an explicit form for the entries $m_{i,j}^{(n)}$ of M^n. Let $a = 2p(1-\lambda)+q$ and $b = 2p\lambda$. Then

$$\begin{cases} m_{k,k}^{(n)} &= a^n(1-s)^{nk} \\ \\ m_{k,k+j}^{(n)} &= a^{n-j}b^j(1-s)^{[j(j-1)/2+nk]} \displaystyle\prod_{l=1}^{j} \frac{(1-s)^{n+1-l}-1}{(1-s)^l-1} \end{cases} \tag{1}$$

Note that for $j = 0$, the formulas do coincide but we state $m_{k,k}^{(n)}$ separately for convenience.

0.3 Extinction

The pgf of the number of mutation-free offspring of a mutation-free individual is

$$\varphi_{00}(t) = \varphi_0(1 - q_0 + tq_0).$$

Now let T be the time until extinction of the mutation-free class, starting from one mutation-free ancestor. Then

$$\{T = n\} = \{Z_n = 0\} \setminus \{Z_{n-1} = 0\}$$

and hence

$$P(\{T = n\}) = P(\{Z_n = 0\}) - P(\{Z_{n-1} = 0\}) = \varphi_{00}^{(n)}(0) - \varphi_{00}^{(n-1)}(0).$$

The mean of T is easiest computed as

$$E[T] = \sum_{n=0}^{\infty} P(T > n) = \sum_{n=0}^{\infty} (1 - \varphi_{00}^{(n)}(0)) \tag{2}$$

where $\varphi_{00}^{(0)}(0) = 0$. The variance is

$$\mathrm{Var}[T] = E[T(T-1)] + E[T] - E^2[T]$$

where

$$E[T(T-1)] = 2\sum_{n=1}^{\infty} nP(T > n) = 2\sum_{n=1}^{\infty} n(1 - \varphi_{00}^{(n)}(0)).$$

If the population is instead started from n_0 mutation-free ancestors, the time until extinction is the time until extinction of the n_0 independent sub-populations started from these ancestors. Let T_k be the time of extinction of the kth ancestor, $k = 1, 2, ..., n_0$ and let T be the time of extinction of the entire mutation-free class. Then

$$T = \max\{T_1, ..., T_{n_0}\}$$

and hence

$$P(T > n) = 1 - P(T \leq n) = 1 - \prod_{k=1}^{n_0} P(T_k \leq n) = 1 - (\varphi_{00}^{(n)}(0))^{n_0}$$

and we obtain

$$P(T = n) = P(T > n-1) - P(T > n) = (\varphi_{00}^{(n)}(0))^{n_0} - (\varphi_{00}^{(n-1)}(0))^{n_0},$$

$$E[T] = \sum_{n=0}^{\infty} (1 - (\varphi_{00}^{(n)}(0))^{n_0}) \tag{3}$$

and

$$E[T(T-1)] = 2 \sum_{n=1}^{\infty} n(1 - (\varphi_{00}^{(n)}(0))^{n_0}).$$

From the previous section, we get that

$$\varphi_{00}(t) = \varphi_0(1 - q_0 + tq_0).$$

Example. If reproduction is Poisson with mean 1, so that the population stays constant on the average but the mutation-free class is sub-critical because of the loss due to mutation, we get

$$\varphi_{00}(t) = e^{1-q_0+tq_0-1} = e^{q_0(t-1)}$$

that is, the number of mutation-free offspring is Poisson with mean q_0. If further the number of mutations is Poisson with mean λ (see Haigh,...),

$$q_0 = e^{-\lambda}$$

and the expression for $\varphi_{00}(s)$ becomes

$$\varphi_{00}(t) = e^{e^{-\lambda}(t-1)}.$$

Example. Binary reproducion. With X the number of offspring $P(X = 2) = 1 - P(X = 0) = p$. Either one or no mutations, mutation probability λ. Then

$$\varphi_0(t) = pt^2 + 1 - p$$

and hence

$$\varphi_{00}(t) = p(1 - \lambda + t\lambda)^2 + 1 - p.$$

0.4 Extinction of Consecutive Classes

Start with n_0 mutation-free individuals and denote by T_0 the time (generation) of extinction of this class. At time T_0, the number of individuals in class 1 (those with one mutation) is therefore

$$m_{0,1}^{(T_0)}$$

which we note is a random variable of the form $g_{0,1}(T_0)$ (function of T_0), where $g_{0,1}$ is given explictly by (1). Thus, the expected number of individuals in class 1 at the time of extinction of class 0 is

$$\begin{aligned} E[m_{0,1}^{(T_0)}] &= E[g_{0,1}(T_0)] \\ &\approx g_{0,1}(E[T_0]) \end{aligned}$$

where $E[T_0]$ is given by (3) and we use a first-order Taylor approximation. To generalize the idea, let $a = 2p(1-\lambda)+q$ and $b = 2p\lambda$ as before, and define the functions

$$\begin{cases} g_{k,k}(x) &= a^x(1-s)^{xk} \\ g_{k,k+j}(x) &= a^{x-j}b^j(1-s)^{[j(j-1)/2+xk]}\displaystyle\prod_{l=1}^{j}\frac{(1-s)^{x+1-l}-1}{(1-s)^l-1} \end{cases} \tag{4}$$

Note that for nonnegative integers n, $g_{k,k+j}(n) = m_{k,k+j}^{(n)}$ for $j \geq 0$, and that the two epressions conincide for $j = 0$. Thus, we have the approximation

$$E[m_{k,k+j}^{(T_k)}] \approx g_{k,k+j}(t_k)$$

the expected number of $(k+j)$-types at the extinction of type k.

We can now establish recursive formulas for consecutive extinction times. Thus, let T_0 be the extinction time for type 0, T_1 the *additional* time until extinction of type 1, and so on. Let $t_k = E[T_k]$. The extinction time of type k is then

$$\tau_k = T_0 + T_1 + \cdots + T_k$$

with mean

$$E[\tau_k] = t_0 + t_1 + \cdots + t_k$$

The extinction time of the entire population is

$$\tau = \lim_{k\to\infty} \tau_k$$

and by monotone convergence this has mean

$$E[\tau] = \lim_{k\to\infty} E[\tau_k]$$

By (2) the consecutive expected extinction times can be computed as

$$t_k = 1 + \sum_{n=1}^{\infty} \left[1 - (\varphi_{kk}^{(n)}(0))^{n_{k-1}^{(k)}}\right]$$

where

$$n_{k-1}^{(k)} = n_{k-1}^{(k-1)} g_{k-1,k}(t_{k-1}) + n_{k-2}^{(k)} g_{k,k}(t_{k-1})$$

and

$$\varphi_{k,k}(t) = 1-(p+q)w(k)+p(1-s)^k\lambda^2+t(1-s)^k(2p\lambda(1-\lambda)+q)+t^2p(1-s)^k(1-\lambda)^2$$

0.5 The size of the fittest class

Anohter question of interest is the relative size of the fittest class is at the time of extinction of the thitherto fittest class. Recall t_0, the expected time of extinction of the 0-class. We can use the function g introduced in (4) to approximate the relative expected size of the 1-class as

$$r_1(t_0) = \frac{g_{0,1}(t_0)}{\sum_{j=1}^{t_0} g_{0,j}(t_0)}$$

Note that at time t_0, the highest possible class present in the population is t_0. Generally, at time t_k, the relative size of class $k+1$ is

$$r_{k+1}(t_k) = \frac{g_{k,k+1}(t_k)}{\sum_{j=1}^{t_k} g_{k,k+j}(t_k)}$$

where the t_k are computed according to (2).

0.6 IMON

We're interested in two things:

(1) the conscecutive mean extinction times $t_0, t_1, ...$ as well as $[\tau_n]$ and $E[\tau]$ in the notation from previous sections

(2) the consecutive relative sizes of the fittest class.

The input parameters are n_0, p, q, and λ. Let $a = 2p(1-\lambda)+q$ and $b = 2p\lambda$. The parameters must be chosen such that $a < 1$ (otherwise theory tells us that $t_0 = \infty$). The total number of offspring is $a+b$ and this **can** be greater than 1; in fact, that might be an interseting case. First compute t_0. We have

$$t_0 = 1 + \sum_{n=1}^{\infty} \left[1 - (\varphi_{0,0}^{(n)}(0))^{n_0}\right]$$

where

$$\varphi_{0,0}(t) = 1 - (p+q) + p\lambda^2 + t(2p\lambda(1-\lambda)+q) + t^2 p(1-\lambda)^2$$

and the $\varphi_{0,0}^{(n)}(0)$ are computed recursively by the relation

$$\varphi_{0,0}^{(n)}(0) = \varphi_{0,0}(\varphi_{0,0}^{(n-1)}(0))$$

Next compupte $n_0^{(1)}$, the expected number of 1-types at the time of extinction of type 0, by using the functions

$$\begin{cases} g_{k,k}(x) &= a^x(1-s)^{x(k-1)} \\ \\ g_{k,k+j}(x) &= a^{x-j}b^j(1-s)^{[j(j-1)/2+x(k-1)]}\prod_{l=1}^{j}\dfrac{(1-s)^{x+1-l}-1}{(1-s)^l-1} \end{cases} \tag{5}$$

for $k=1$. Specifically, we have

$$n_0^{(1)} = n_0 g_{0,1}(t_0)$$

The expected numbers in the other classes are

$$n_0^{(j)} = n_0 g_{0,j}(t_0)$$

for $j = 1, 2, ..., t_0$. The relative size of class 1 is therefore

$$r_1(t_0) = \frac{g_{0,1}(t_0)}{\sum_{j=1}^{t_0} g_{0,j}(t_0)}$$

Now proceed to computing t_1, and so on. The general formula is

$$t_k = 1 + \sum_{n=1}^{\infty}\left[1 - (\varphi_{k,k}^{(n)}(0))^{n_{k-1}^{(k)}}\right]$$

where

$$n_{k-1}^{(k)} = n_{k-1}^{(k-1)} g_{k-1,k}(t_{k-1}) + n_{k-2}^{(k)} g_{k,k}(t_{k-1})$$

and

$$\varphi_{k,k}(t) = 1-(p+q)w(k)+p(1-s)^k\lambda^2+t(1-s)^k(2p\lambda(1-\lambda)+q)+t^2p(1-s)^k(1-\lambda)^2$$

The relative size of class $k+1$ is

$$r_{k+1}(t_k) = \frac{g_{k,k+1}(t_k)}{\sum_{j=1}^{t_k} g_{k,k+j}(t_k)}$$

The t_k should eventually go to 0 quickly. When they are below 1, this means that the population is exxpected to go extinct within less than one generation, which in practice means that they are very likely to go extinct soon.