

The Probability of Escaping Accidental Extinction

Alessandro Giusti

September 17, 2006

Abstract

This report is about the proof for theorem (C.1), which allows to compute the probability that a mutation – initially appearing in a single individual – escapes accidental extinction, in the framework of the Adaptive Dynamics.

1 Introduction

The Adaptive Dynamics (AD) approach allows to study evolutionary dynamics of traits in competing populations, providing a system of differential equations, each driving the evolution of one trait.

This result is achieved after several simplifying assumptions: a simple asexual, physiologically unstructured demographic model (so that genetics must not be taken into consideration), homogeneous environments, and continuous traits (real dimensionless variables) mutating rarely, independently and with infinitesimal variations.

The main premise of the AD approach is that evolution originates from mutation and competition processes. Under the previous assumptions, when a mutation happens a single mutated individual appears when resident populations are at equilibrium. If the mutation is advantageous, the *invasion implies substitution theorem* shows that the mutant population replaces the corresponding resident population; the smooth evolution of traits – which develops in a much broader timescale – is a consequence of continuous successful mutations and substitutions.

However, not all advantageous mutations succeed: the mutant population only thrives if it escapes accidental extinction; in order to derive the AD canonical equation, the probability of accidental extinction must be quantified.

1.1 Why the mutant population should extinguish?

The possibility of accidental extinction of the newborn mutant population can not be explained by the resident-mutant model, if the mutation improves the fitness; on the contrary, one would expect that a mutant population bearing an advantageous trait would necessarily thrive and substitute the corresponding resident population.

This apparent contradiction is solved by noting that the resident-mutant model (and the invasion implies substitution theorem) do not hold when the number of individuals of the mutant population is small. This happens because the resident-mutant model uses a continuous real variable for population abundance and abstracts the stochastic birth and death processes to a deterministic equation – which is justified by the large number of individuals involved; but if the individuals of a population are few, the underlying stochastic processes must be considered directly. This need arises immediately after each mutation happens: in fact, a *single* mutated individual appears. Even if the mutated trait grants a slightly higher per-capita growth rate to the newborn mutated population, the single mutated individual is still subject to a probability of death for each time interval; as soon as the number of mutated individuals reaches 0, there is no way out and the mutant population faces accidental extinction. If a new mutated individual is given birth before the first one dies, extinction is still possible since there is still a (lower) probability that both die before a new one is generated; accidental extinction can not be ignored until the mutant population is composed by a sufficiently large number of individuals.

Note that, similarly, after a disadvantageous mutation the new population faces extinction without reaching a considerable number of individuals; in this case, however, extinction would not be “accidental” since it is justified by the resident-mutant model as well. There is no such thing as an “accidental thriving” since, should the new population reach by an unlikely chance a sufficiently large number of individuals to be accurately modeled by the resident-mutant model, the large-scale dynamics of the resident-mutant model would lead it to extinction nonetheless.

2 Generalities

The proof works by considering the Markov chains explaining the evolution of the mutant population; birth and death are considered independent Markov processes. The per-capita birth and death rates (λ_b and λ_d) are independent on the abundance of the mutant population and are consid-

ered constant: in fact, during this analysis we keep close to the equilibrium of the resident-mutant model, at which the resident population abundances are constant, because the mutant population is of minimal size w.r.t resident populations.

Each node in the Markov chain is the number of individuals of the mutant population, and under the presented assumption this is a complete description of the state (i.e. the transition probabilities are not defined by anything else). The initial state is 1, i.e. a mutant individual has just appeared. 0 is an absorbing state, because there are no outbound transitions – the mutant population is extinguished.

The probability that more than one birth and/or death event occurs during the same time interval is 0 because we are considering an infinitesimal dt : therefore we will only have transitions connecting a state I to state I , $I - 1$ and $I + 1$.

Because we are dealing with *per capita* rates, with only one mutant individual the probability of a birth in the considered time interval is $\lambda_b dt$; in state I , the probability of a birth event occurring is $I\lambda_b dt$, whereas the probability of a death is $I\lambda_d dt$.

This explains the markov chain in figure C.1.

2.1 Goal

What we want to prove is (C.1), where $P_{1,I}$ is the probability that a mutant population initially composed by a single individual is after some time composed of I individuals. This is significant because by taking the limit for $I \rightarrow \infty$ allows to find (3.35), which expresses the probability that a mutant population initially composed by a single individual escapes accidental extinction. As we anticipated, that limit is obviously greater than 0 only if the mutation increases the fitness of the new population (which is expressed after the “if” in (3.35)), which, by the way, is one of the preconditions (3.23) for the invasion implies substitution theorem.

3 Proof

Our thesis can be rewritten, by means of (C.4) – a well known result about geometric series – as (C.5), where $q = \lambda_d/\lambda_b$; (C.5) is then proven by induction; the base case is obviously met, since $P_{1,1} = 1$ and $(q^0)^{-1} = 1$.

Therefore, the rest of the proof aims at proving the induction step, i.e. (C.6) giving (C.5) as granted. In order to do that, an expression for $P_{1,I+1}$ is

derived by doing some transformations on Markov chains and by exploiting (C.5) more than once, eventually proving (C.6).

3.1 Transformations on the Markov Chains

In order to understand how the induction step is proven, a number of transformations of the markov chain in figure C.1 are introduced.

The markov chain in figure C.1 is not convenient since it has an infinite number of states; however, if our goal is only to compute $P_{1,I}$, we can substitute chain C.1 with chain C.2, “truncating” chain C.1 to state I which becomes absorbing: in fact, as soon as the number of individuals reaches I in a realization, we have determined that the realization will contribute to $P_{1,I}$, regardless on what happens next.

Chain C.3 depicts the same transformation, this time for computing $P_{1,I+1}$.

Chain C.4 introduces a meaningful simplification of chain C.3, in which only 5 states ($0, 1, I-1, I$ and $I+1$) are kept. In chain C.4 the meaning of some of the transitions is not straightforward; for example, $1 - P_{1,I}$ accounts for all realizations starting at 1, in which state I is not reached in finite time – then, they will eventually reach 0, since that’s the only available absorbing state. Similarly, $P_{I-1,0}$ is the probability that a realization starting at $I-1$ reaches 0, *without ever passing at I* . This clause is important because it leads to the most tricky step of the proof.

3.2 Deriving $P_{I-1,0}$

$P_{I-1,0}$ can be derived as the probability of reaching state 0 from state $(I-1)$ in chain C.2: in fact, all realizations which reach 0 from $I-1$ in chain C.2 do not pass from I since it is an absorbing state.

The interesting part is that chain C.2 exhibits a simmetry which can be exploited in order to compute $P_{I-1,0}$. In fact, by renaming state I to “0’”, state $I-1$ to “1’”, and so on until all states are renamed (state 0 becomes state “I’”), you obtain an identical chain in which λ_b and λ_d are swapped. Therefore computing $P_{I-1,0}$ in C.2 is the same as computing $P_{1,I'}$ in the renamed chain. Since (C.5) holds (inductive hypotesis), we obtain (C.7) – note that q is now $1/q$ since λ_b and λ_d are swapped; the second equality in (C.7) is only an algebraic step.

This derivation of $P_{I-1,0}$ is useful when deriving $P_{1,I+1}$.

3.3 Deriving $P_{1,I+1}$

Our goal is to find an expression for $P_{1,I+1}$: this is possible by means of chain C.5, which is a transformation of chain C.4, removing state $I - 1$. Edge $(I,0)$ in chain C.5 is the result of probability product and is computed easily; edge (I,I) in chain C.5 is computed by summing the probabilities of realizations $(I,I - 1,I)$ and (I,I) in chain C.4.

$P_{1,I+1}$ is computed as the sum of the probabilities of all realizations going from state 1 to state $I + 1$ in chain C.5: the only possible branch is at state I , since the edge (I,I) can be followed 0 or more times; by using the same geometric series as in (C.4), (C.8) is obtained.

Other than λ_b , λ_d , dt and I , in all transition probabilities of chain C.5 only appear terms $P_{1,I}$ and $P_{I-1,0}$, which are defined in (C.5) and (C.7); by merging (C.5), (C.7), (C.9) into (C.8) all the variables blend nicely, and give (C.6), which proves the theorem.