Analysis of Fixation of Beneficial Mutations in Populations of Variable Size

Ozge I. Ozel
Department of Physics,
Massachusetts Institute of Technology

(Dated: May 12, 2011)

Beneficial mutations occurring within reproducing populations are at the heart of many factors, which determine the fate of a population after many generations. This paper presents the results of a study in which the fixation probability of a beneficial mutant allele is analyzed under different demographic conditions, namely single change, exponential and logistic growth (decline). The conclusion is that the fixation probability is highly dependent on changes in population size.

PACS numbers:

I. INTRODUCTION

Across generations the ability of individuals to adapt within a population crucially depends on the fixation of beneficial mutations. The pioneering studies of Fisher[1], Haldane[2] and Wright[3] in the early 30s have indeed shown that the probability that a beneficial mutation will fix within a large population can be approximated by $2s$, where s is the selective advantage of the mutant allele. This seemingly low value confirms clearly the stochastic nature of genetic variations in individuals within large populations.

In their analyses, Fisher, Haldane and Wright made several assumptions, one of which is that the size of the population is constant, which will be addressed in this paper by changing the size under certain demographic models. In this conjecture, families are assumed to have a Poisson distribution of reproductive success and any mutant allele starts out as a single copy. Kimura[4] was the first to describe the behavior of models of this sort by a method called diffusion approximation to prove that the fixation probability of a mutant allele with initial frequency p and selective advantage s is

$$\frac{1-e^{-4N_e sp}}{1-e^{-4N_e s}}$$

with N_e being the variance effective size of a diploid population. This formulation is applicable to arbitrary values of p and is even valid for other distributions of reproductive success as long as the size of the population remains around the same average.

To address this major restriction of the model, so that the fixation of a mutant allele could be analyzed in populations of varying size, which is usually the case in nature, a numerical analysis was performed by Kojima and Kelleher[5]. And not long after, Ewens[6] shows that the fixation probability is approximately equal to $2sN'/N_t$, where N' is the harmonic mean of the population sizes N_t in each generation.

In this paper, we adapt part of the formulation developed by Otto and Whitlock[7] to analyze the effect of changing population size on the fixation probability of beneficial mutant alleles. The conditions under which the assumptions apply are growth (decline) in a single change, exponentially or logistically. This analysis can further be extended to insertion of deleterious mutations within changing populations, which will however not be covered in this current discussion.

II. FIXATION PROBABILITY FOR A POPULATION WITH CONSTANT SIZE

In his study, Haldane uses a Poisson progeny distribution in a branching process, which ensures that the expected number of offspring is one, meaning that the overall population size remains approximately constant. When a beneficial mutation occurs in the population, it will possess higher viability over the wild-type resulting in $s+1$ times more progeny, where s is a measure of the relative fitness advantage of a heterozygote in a randomly mating diploid population. If the probability of fixation for a mutant allele, that appears in generation t is given by P_t, $1-P_t$ gives the probability for loss of that allele. The probability that all k alleles will result in no offspring, $(1-P_{t+1})^k$, times the probability of having i offspring gives the probability that each of its offspring fail to leave offspring after a long time, when summed over all possible values of k.

$$\begin{aligned} 1-P_t &= \sum_{k=0}^{\infty} \frac{(s+1)^k}{k!} e^{-(s+1)} (1-P_{t+1})^k \\ &= \exp\left(-(s+1)P_{t+1}\right) \qquad (1) \end{aligned}$$

The fixation probability at generation t is equal to that of the next generation, $P_t = P_{t+1}$, when s and the population size are assumed to be constant. Hence, the solution for s can be evaluated in a straightforward manner, where P_t^* denotes the fixation probability under the above stated condition.

$$s = -\frac{\ln(1-P_t^*)}{P_t^*} - 1 \qquad (2)$$

If s is small, P_t^* is also small, hence ignoring second and higher order terms, the expansion of Eq. (2) would lead to the fixation probability that is twice the relative fitness, $P_t^* \approx 2s$.

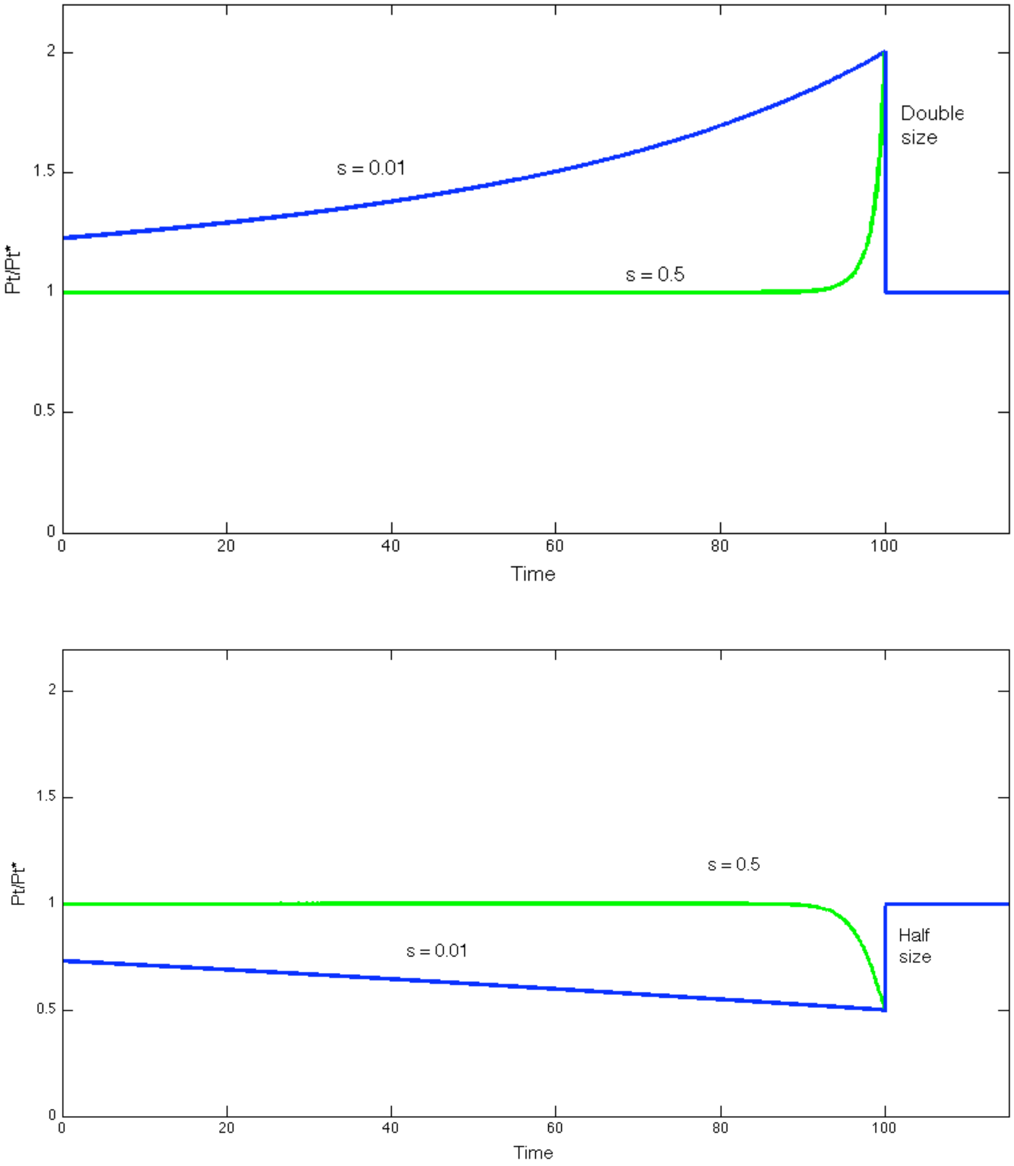

FIG. 1. Changes in fixation probability of beneficial alleles in 100 generations when population size is doubling (top) and halving (bottom) in a single change. The fixation probability is shown relative to the fixation probability of a population of constant size.

III. FIXATION PROBABILITY FOR A POPULATION WITH VARIABLE SIZE

For a population varying in size, we can adopt Eq. (1) using a discrete generation model as follows. Let N_t denote the population size at generation t and N_{t+1} the size at generation $t+1$, so that ΔN_t equals the change from generation t to $t+1$. The average number of descendants left by each individual would then be N_{t+1}/N_t. Under the same assumptions, which were valid for the constant population size case, we can now modify Eq. (1)

$$\begin{aligned} 1 - P_t &= e^{-(s+1)\dfrac{N_{t+1}}{N_t}P_{t+1}} \\ &= e^{-(s+1)\left(1+\dfrac{\Delta N_t}{N_t}\right)P_{t+1}} \end{aligned} \tag{3}$$

as was demonstrated by Ewens[6]. This expression shows that for a growing population the fixation probability for beneficial mutations is always higher than $2s$, while on the contrary for shrinking populations individuals, who carry the beneficial allele, have a tendency to fail to have offspring. In a growing population, the fixation probability is higher and the mutant allele is less likely to be lost because it is rare in the population and hence more likely to be established.

In the next three subsections, we will analyze three different types of growth models, namely single change in population size, the exponential and the logistic growth to analyze the effect on the probability of fixation. The calculations are performed using Eq. (3) and its continuous time approximation

$$\begin{aligned} 1 - P_t = 1 - (s+1)\left(1 + \frac{\Delta N_t}{N_t}\right) P_{t+1} \\ + \frac{(s+1)^2}{2}\left(1 + \frac{\Delta N_t}{N_t}\right)^2 P_{t+1}^2 + \cdots \end{aligned}$$

$$\begin{aligned} P_t - P_{t+1} &\approxeq \frac{\Delta N_t}{N_t} P_{t+1} + sP_{t+1} - \frac{1}{2}P_{t+1}^2 \\ \frac{dP_t}{dt} &= -\left(s + \frac{1}{N_t}\frac{dN_t}{dt}\right) P_t + \frac{1}{2}P_t^2 \end{aligned} \tag{4}$$

where we made the assumption that s, P_t, and $(1/N_t)\frac{dN_t}{dt}$ are small and have ignored terms higher than second order.

A. Change in population size for a single generation

Let us consider the case when the population size is initially S_0 in generation T and changes to S_1 in generation $T+1$ but remains constant at all other times. Also the probability of fixations after generation $T+1$ remains the same, P_t^* as used before. In setting $P_{T+1} = P_t^*$ in Eq. (3) with s and P_t^* assumed to be small, we get $P_T = S_1 P_t^*/S_0$ for the fixation probability at generation T. Now, using this expression and solving the differential equation given in Eq. (4), we arrive at

$$P_t \approxeq \frac{S_1}{S_1 + (S_0 - S_1)e^{-s(T-t)}} \tag{5}$$

where $dNt/dt = 0$. This result implies that when s is small, changes in the size of the population can affect the fate of mutations that arose many generations earlier. The ratios of the fixation probabilities are plotted in Fig. 1 for $s = 0.001$ and $s = 0.5$. Beneficial mutations with weak selective advantage are comparably more affected when the size of the population changes only once. This single change in size backpropagates to modify the fixation probabilities of previous generations.

B. Exponentially growing (declining) population

When the population grows with constant rate a, which is taken to be small, part of the exponent in Eq. (3) is replaced by $(s+1)(a+1)$ instead of $(s+1)$. Keeping only terms up to first order in s and a, we get $s+a+1$. From this, we find that for a continuously growing (shrinking) population, the probability of fixation is approximately $P_t \approx 2(s+a)$, when $s+a>0$ and s, a are sufficiently small. If $s+a<0$, increase in the size of the population will eventually lead to the loss of weakly selected mutant alleles, which imposes a sensitive dependence for the fixation probability on the growth rate.

C. Logistically growing (declining) population

The size of a population that is growing logistically is bounded by its carrying capacity that critically depends on conditions of the particular environment. The rate of population increase is limited by the population density, and hence decreases linearly with the size of the population. This model is described by the following equation

$$\frac{dN_t}{dt} = aN_t\left(1 - \frac{N_t}{C}\right) \tag{6}$$

whose solution is

$$N_t = \frac{S_0 C}{S_0 + e^{-at}(C - S_0)} \tag{7}$$

where C is the carrying capacity, a the intrinsic growth rate and S_0 the initial size of the population. Now, substitution of Eq. (6) into Eq. (4) would yield

$$\begin{aligned} \frac{dP_t}{dt} &= \frac{dP_t}{dN_t}\frac{dN_t}{dt} \\ &= -P_t\left(s + a\left(1 - \frac{N_t}{C}\right)\right) + \frac{P_t^2}{2} \end{aligned} \tag{8}$$

where we have used the chain rule in the upper line. This results then in the following expression

$$\frac{dP_t}{dN_t} = -\frac{P_t\,(sC + a(C - N_0))}{aN_t(C - N_t)} + \frac{P_t^2 C}{2aN_t(C - N_t)} \tag{9}$$

with its solution straightforwardly calculated as

$$P_t \approxeq \frac{2(s+a)sC}{sC + aN_t} \tag{10}$$

From Eq. (10) we can obtain our previous result for the fixation probability of a population of constant size $N_t = C$ as $2s$ and $2(s+a)$ for an exponentially growing population for a population size which should be much smaller than the carrying capacity. Substituting the expression in Eq. (7) into Eq. (10) gives us the dependence of the fixation probability with respect to time.

$$P_t = \frac{2s(s+a)(S_0 + e^{-at}(C - S_0))}{S_0(s+a) + se^{-at}(C - S_0)} \tag{11}$$

Numerical estimates have shown that the diffusion approach studied by Kimura provides more accurate results

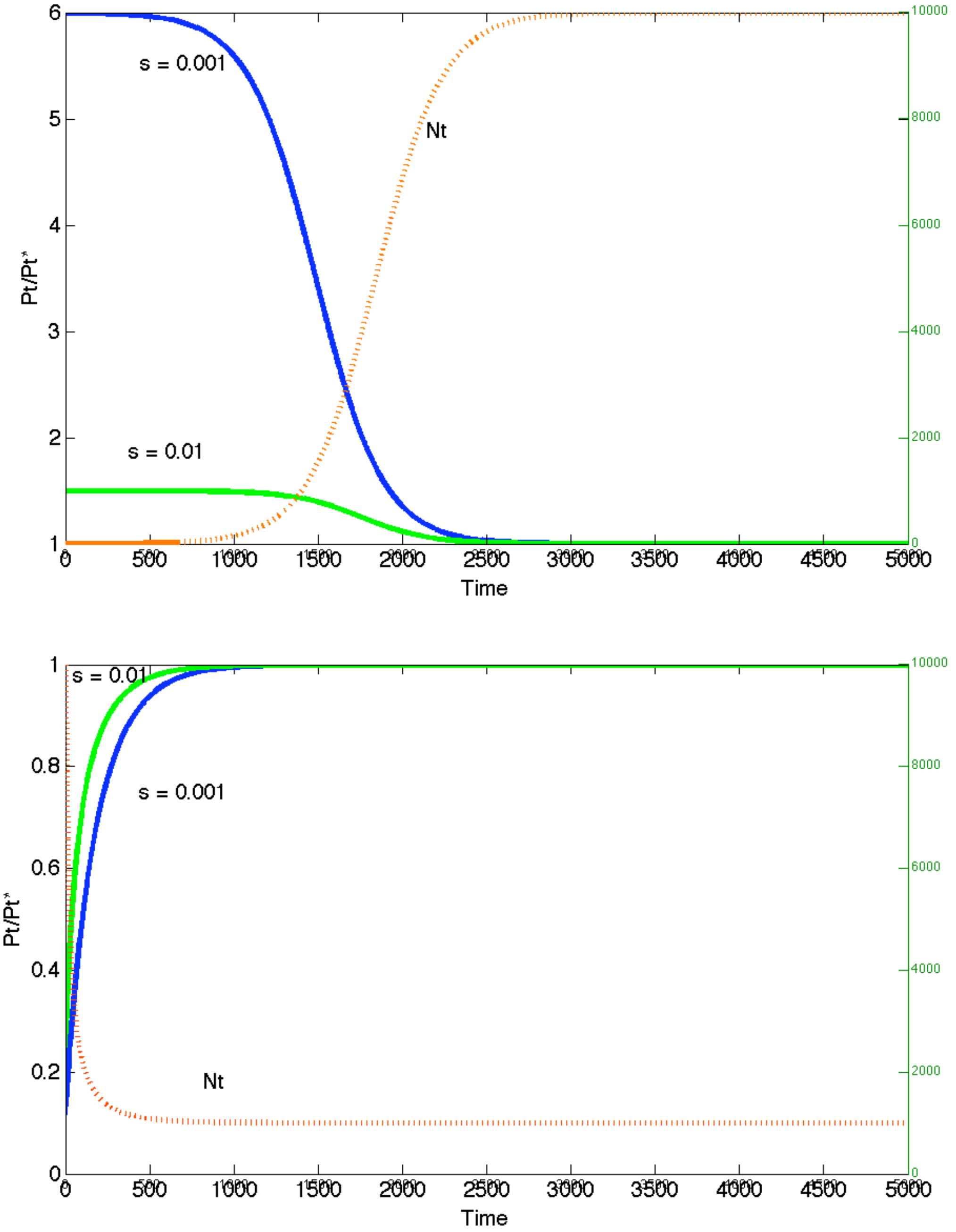

FIG. 2. Changes in fixation probabilitysolid lines, and population sizedotted line, of the beneficial alleles under the logistic model when the population size starts off from 1 with carrying capacity 10,000 (top) and size is initially 10,000 with carrying capacity 1,000 (bottom). The fixation probability is shown relative to the fixation probability of a population of constant size. The growth rate is $a = 0.005$. A pronounced change in the fixation probability is observed in early generations when the growth rate is nearly exponential and decreases when the population approaches the carrying capacity.

when either s or a are large or the size of the population is small.

In Figure 2, we have plotted changes in the fixation probability for two values of selective advantage, s starting from a population size below and above the carrying capacity. In both cases, the probability of fixation approaches its limiting value $2s$ as the size of the population reaches its carrying capacity. Also, an important feature observed in early phases of exponential growth, is the prominent change in the fixation probability, which disappears when the carrying capacity is reached. When there is weak selection and the initial size of the population is far from the carrying capacity, changes in the size substantially affect the probability of fixation.

IV. CONCLUSIONS

The current analysis of the fixation probability for beneficial mutations on three different growth models has shown the strong dependence of on changes in population size. Our results have shown that within a growing population, selection has an enhanced effect, meaning that the probability that a beneficial mutant allele becomes fixed is highly likely. Opposite to what is observed in growing populations, in shrinking populations, selection has a lesser effect, making the beneficial mutation less likely to fix. We have limited our discussion only to the analysis of beneficial mutations, however similar arguments could further be developed to include effects of deleterious mutations on the fixation probability within population cycles, which is extensively discussed in the original paper of Otto and Whitlock[7].

---

[1] R. A. Fisher, Proc. Roy. Soc. Edinb. **42**, 321 (1922).

[2] J. B. S. Haldane, Proc. Camb. Phil. Soc. **23**, 838 (1927).

[3] S. Wright, Genetics **16**, 97 (1931).

[4] M. Kimura, Ann. Math. Statist. **28**, 882 (1957).

[5] I. C. Kojima and T. M. Kelleher, Am. Nat. **96**, 329 (1962).

[6] W. J. Ewens, Heredity **22**, 438 (1967).

[7] S. P. Otto and M. C. Whitlock, Genetics **146**, 723 (1997).