1 Introduction

Evolution in populations has long been a topic of interest in the community of biologists and ecologists.[1] Such evolution can be observed in any group of organism that tightly or loosely form a population that interact with one another, from bacteria colonies to animals and human communities. In such population evolution, it is often assumed that some mutation occurs in a random individual with some given probability, and this mutation could be either beneficial, neutral, or deleterious. After some number of reproduction cycle, this mutation that has appeared would eventually either go to fixation, which means it has spread through the entire population, or extinction, which means it has disappeared from the population.[5] Often, we are interested in the probability of fixation, which is the probability that the mutant would eventually fix in the population. Besides this, time to fixation is also an important metric that is often explored.

Previous studies have established the correlation between selective advantage and probability of fixation. Intuitively, with a higher selective advantage that directly contributes to reproductive fitness, there is a higher probability of fixation. Previous analytical approaches have concluded that with a neutral mutation, the probability of fixation can be well-approximated by $1/N$ [6]. However, this assumes a well-mixed population, in which every individual is allowed to interact with every other individual, whereas in reality, this is hardly the case, and populations are often posited on a less-dense network, which is inherently different from a well-mixed network, and could lead to different results when it comes to evolution.

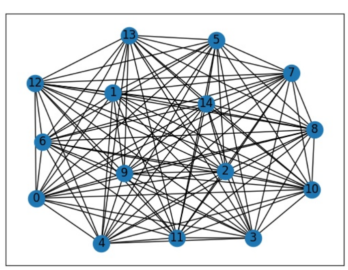

Figure 1: Illustration of a well-mixed population, edges represent an evolutionary corridor – an allowed birth-death replacement link

Previous work by Y.P. Kuo has established that populations on a structured network responds differently to selective advantage of mutations.[5, 3] For instance, a star graph, which is very sparse and have highly variable degree distribution, responds to selective advantage s differently from a well-mixed network. Specifically, the star graph has been found to be an amplifier, which amplifies the selection advantage of the mutant.

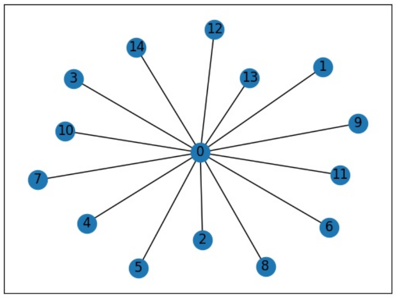

Figure 2: Illustration of a population on a star graph, edges represent an evolutionary corridor – an allowed birth-death replacement link

However, it should be noted that in nature and almost all biological settings, the mutant doesn't always contribute to the same selective advantage at all times. Often, there are environmental factors that either slightly or significantly modifies the selective advantage of the mutant.

To date, no literature work has been done on how such variability could affect the probability of fixation in a structured population. We believe that through this study, we are filling in a gap in which the selective advantage, and consequently fitness in a structured population are no longer fixed to be a constant, but rather have the same mean but for a range of variance. Here, we are specifically interested in the case when environmental factors contribute slightly to selective advantage and relative fitness, with such environmental factors represented as noise or variance in the mutant population. Through this project, we aim to answer the question of how variance changes the probability of fixation, in the context of different network structure and in the setting of various selective advantage.

2 Approach

This project did not start off with the goal of investigating the impact of fitness in mind. In fact, this project started off with designing communication protocols which in theory should lead to a rugged fitness landscape. The idea is that with a not so effective or contradictory communication pattern, the fitness of a population should decrease, and conversely, with a highly effective and accurate communication protocol, the fitness of the population should be higher. It is during the process of playing with such fitness landscape that we discovered the heavy presence of noise in fitness, and moved on to studying the implication of such noises, which make up the variance in fitness distribution.

2.1 Exploration in communication protocols and fitness valley

As explained, my initial idea was to create a fitness valley such that a population could move towards a better communication protocol by going through a deleterious intermediate. In figure 3 below, the x-axis is a communication truthfulness parameter, in which -1 means it always lies, and 1 means it always tells the truth. I was expecting to see a clear concave curve, as the fitness of the population moves down and then up as communication becomes more accurate. However, the graph appears to be extremely flat, and although there are some tidy differences at the edge, there is no clear trend at all.

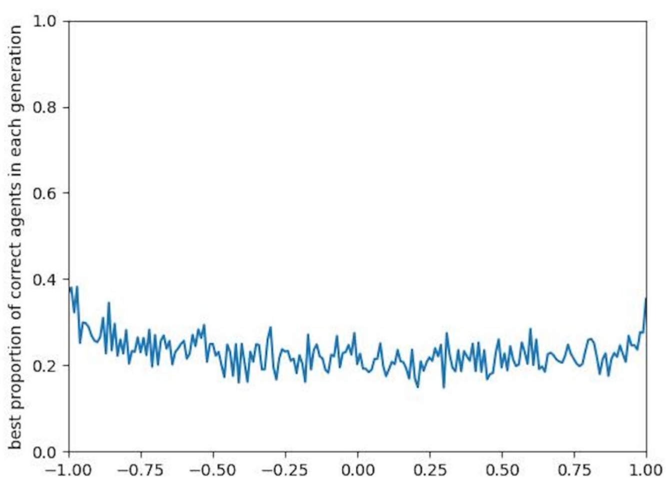

Figure 3: Fitness with different parameter regarding truthfulness in reporting accuracy

In the process of running and observing thousands of simulations, I realized that there is a lot of stochasticity in the initial condition and the Brownian motion afterwards, which could contribute to the lack of clear trend and the lack of smoothness in the curve. Since our eventual goal is to run evolution on a structured population, this observation of noise made us wonder what happens when the fitness are not always fixed but is drawn from a distribution with the same mean but non-zero variance.

2.2 Investigation on distribution of "fitness"

To get some idea of how the "noise" may look like in our simulations, the probability density function of the "fitness", in this case being the fraction of agent reaching the correct site, was plotted. Figure 4 below was the distribution when the mean is around 0.6. In this case, the sample variance was about 0.0025, giving us a standard deviation about 0.05, which is about 8% of the mean. This mean is expected from our experimental set up by a simple analytical calculation.

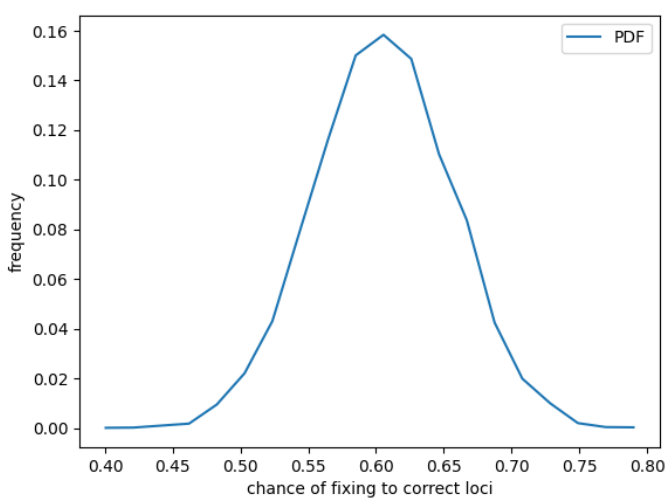

Figure 4: Distribution of a random walk with probability of converging to correct loci 0.6

This gives us some idea about the kind of noise and variance that we are expecting from simulation. With this information, we could approximate the distribution in our further analysis by a random variable instead of having to run multiple simulations.

2.3 Investigation on effect of Variance

The final part of the project is to investigate the effect of variance on probability of fixation on different network structure. Specifically, evolution was performed on a population that is 1) well-mixed, and 2) on a star graph, with the same mean in each case, and the mutant could take on different selective advantage s ranging from 0 to 0.1.

3 Methodology

3.1 Model set-up

To understand the effect of variance, we did multiple(at least 50,000,000 for each parameter) simulations with same mean and different mutant variance. In addition, this is done on both well-mixed and star graph so that we can have some idea about whether variance has different effect on different network structures.

Each simulation follows a Moran birth-death process. At every time step, one of the node on the network gets selected by a probability proportional to its fitness, and it uniformly randomly replaces one of its neighbors. This process continues until either the mutant propagates through the entire population, ie. the number of wild-type individual goes to 0, or the mutant goes extinct, ie. the number of mutant goes to 0.

For the simulations done, the networks used are of size 100, and every time the mutant gives birth to a new individual, the fitness of the new mutant individual is sampled from the specified distribution. In my experiments, uniform and Bernoulli distributions are used for their ease of sampling and ease of analyzing variance.
The code for the simulations are available at:
`https://github.com/ViolaChenYT/RobotSwarmSimulation/tree/main/graph_variance`

3.2 Experimental Evaluation and Results

Through my simulations, I found that variance did have an effect on probability of fixation. Although such an effect is not obvious when the mutation is neutral, there is an effect as shown in figure 6, and such effect is more apparent when the selective advantage of the mutant is larger.

Furthermore, this effect of variance is observed in both well-mixed population and population on a star-shaped network, with the effect being greater in star graphs. In addition, the result so far seems to show that the amplifier properties for graphs, and the corresponding amplification parameters, found in [3], seem to be preserved in the presence of variance.

3.2.1 Selection advantage amplifies effect of variance

With a neutral mutant, ie. selective advantage = 0, it was not clear from figure 5 that variance makes any difference in the fixation probabilities, even if the variance gets as large as it could get, in this case 25. This is not surprising, and has in fact been found by my advisor in [2] earlier, that the effect of variance is very weak. However, the effect of variance seems to be amplified with even a very little selective advantage.

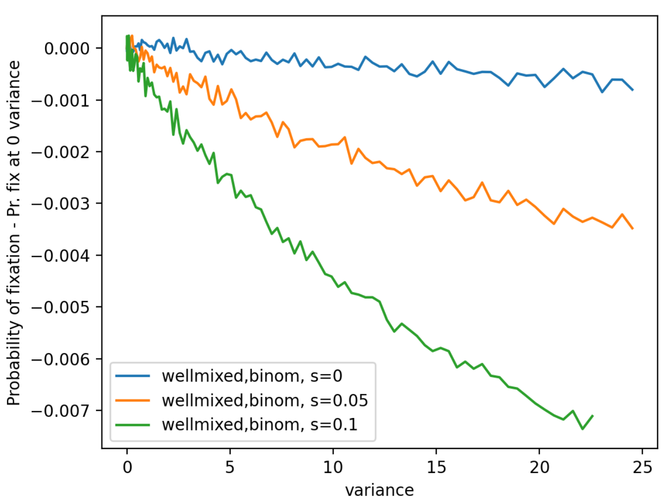

Figure 5: Effect of variance on probability of fixation, under different selective advantage s, y-axis is Prob. fix − Prob. fix at variance 0 with the same selective advantage. Thus, all curves should star at 0 and the rest of the curvatures should reflect only the effect of variance.

Interestingly, although at $s = 0$ we are not able to see much difference in probability of fixation as variance varies, we see much more significant effects for $s = 0.05$ and $s = 0.1$, even though this is not that high a selective advantage compared to the wild-type. We are still trying to understand intuitive why this may be the case, that a small selective advantage could amplify the effect of variance so much.

3.2.2 Effect of variance more substantial on star graph

To look at how variance may interact with different network structures, we compared the probability of fixation, subtracted that when variance is 0, in well-mixed population and star graph. In figure 6 below, we can see for the first time that variance does contribute lower probability of fixation even at variance 0, although the scale on the y-axis suggests clearly that this is not a big decrease, which explains why we are not able to see it on a larger scale. In addition, it is interesting how although the curves are noisy in both cases, mostly due to the high resolution that we are plotting at, the curve for star-graph is almost always below that of the well-mixed, suggesting that for a neutral mutation, variance in mutant fitness has a more significant effect on the star graph, as compared to on the well-mixed population.

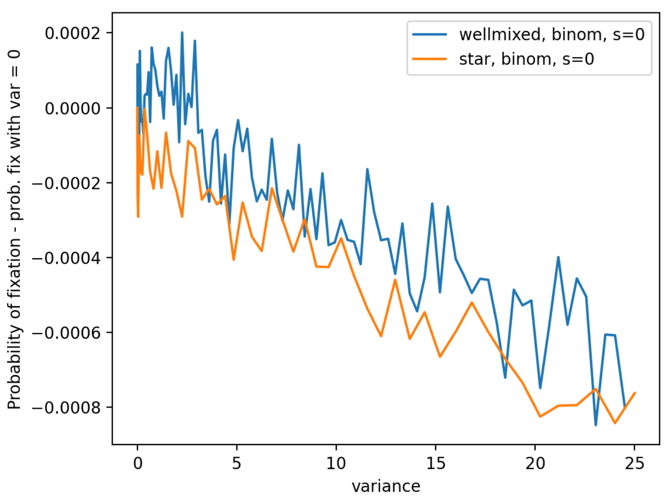

Figure 6: Effect of variance on probability of fixation in well-mixed vs. star graph when selective advantage is 0 ie. no advantage

We then plotted the same graph when the mutant is slightly beneficial, reflected by a selective advantage of $s = 0.05$ in figure 7. Here, we see a much clearer divergence in the two curves, as compared to the case when the mutation is neutral. This strengthened our previous idea of variance having a stronger effect on star-graph.

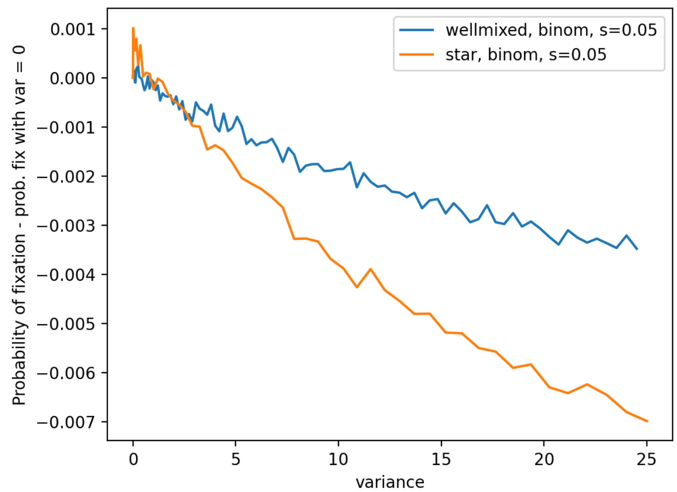

Figure 7: Effect of variance on probability of fixation in well-mixed vs. star graph when selective advantage is 0.05

It should be noted that we observed a crossing of the two curves in this graph at small variance. This seems to suggest that at very small variance, the star graph has a higher probability of fixation compared to no variance, and the effect is lost when the variance gets larger. However, we cannot confirm whether this is purely due to random noises or is this due to subtleties in the network structure. More simulations and work need to be done to either confirm or reject this result.

Part of the reason why we are doubting this phenomenon in figure 7 is that we are not observing a similar trend in figure 8 below, in which the selective advantage of the mutant is just slightly higher. When $s = 0.1$, it seems that the two curves for well-mixed and star graph both start from 0 and diverges as the variance increases, and there is no crossing of the two curves at all. If the trend from figure 7 is real, we would expect a similar trend to be observed here as well.

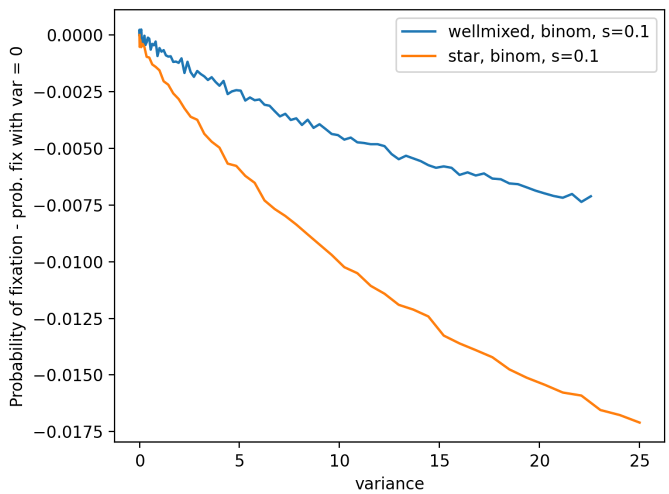

Figure 8: Effect of variance on probability of fixation in well-mixed vs. star graph when selective advantage is 0.1

In the following figure 9, we put together the performance across both types of networks and all selective advantage s. Overall, it is evident that the effect of variance is stronger on star graph, and is stronger with a stronger selective advantage.

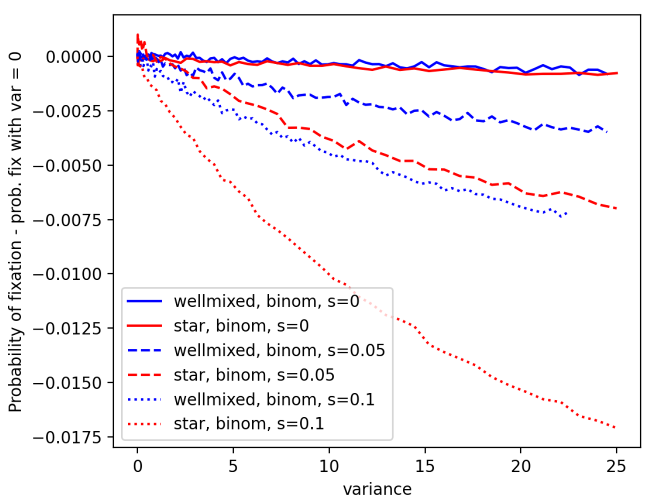

Figure 9: Effect of variance on probability of fixation in well-mixed vs. star graph, with different selective advantage s

3.2.3 Possible relation to previous results on network property

The figure below showed the uncorrected probability of fixation for all network types and selective advantage parameters. Previous results from [3] has found star graph to be am amplifier graph with parameter $\alpha = 2$, which means that it amplifies the effect of selective advantage by a factor of 2.

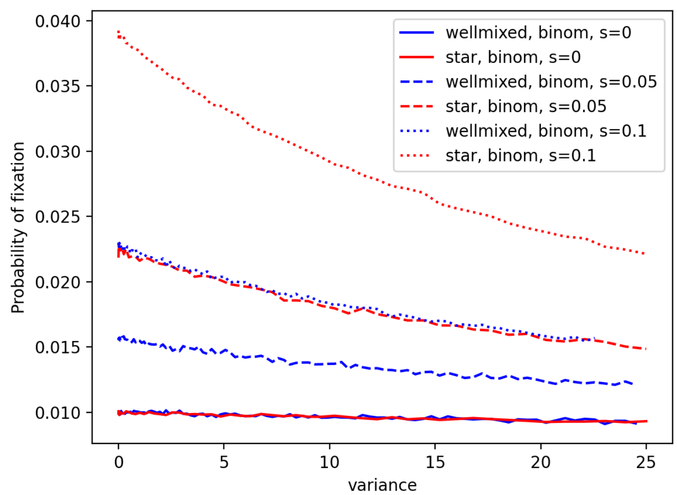

Figure 10: Fitness with different parameter regarding truthfulness in reporting accuracy

Interestingly, in figure 10, we see that the line for well-mixed with $s = 0.05$ almost overlaps with the line for star and $s = 0.1$. We think this is interesting since it seems to correspond to previous finding that star graph should amplify the selective advantage by a factor of two. At the same time, 1) the two curves do not exactly match up, in fact, the two curves for $s = 0$ seem to be more closely together than these two lines. It remains as part of our future work to check if it is true that the amplification factor of networks is preserved in the presence of variance, and what could be a reason for the deviation between the two curves since they do not overlap completely.

4 Surprises and Lessons Learned

4.1 Surprises

1. It was surprising for me to learn that what I thought would make a difference in fitness – the truthfulness parameter mentioned in section 2.1 did not actually have any significant effect.

2. I did not anticipated that the effect of variance when mutation is neutral to be elusive, since I expected that having variable fitness is going to be harmful for the population. What is more surprising is that this effect seems to be amplified even with a small positive selective advantage. We still struggle to understand the reason behind this, but the finding itself is interesting and surprising.

3. I was expecting that with a large enough variance, the geometric mean of the mutant would drop below that of the wild-type, leading to a lower than $1/N$ probability of fixation according the geometric mean principle found in [4].

4.2 Learning points

1. It is common for research to not go as expected. However, what matters is not if it went well the first try, but rather whether we could persevere through and keep trying new things and come up with new ideas until we come up with something that actually works. For this project, I felt hopeless and it seemed like nothing worked for some period of time until we came up with the idea of examining effect of variance.

2. Doing a lot of replications for any experiments or simulations are very important for us to get convincing results and to be able to see trend in the midst of noise.

3. It is important to have good research book-keeping and documentation habits. Even though this is not a big project, there are already a lot of figures and codes and data that need to be handled and managed with care.

5 Conclusions and Future Work

To summarize, we found that in any stochastic fitness simulation, noise, or variance, is almost always created. In such cases, representing fitness with its expected value may not be very accurate. We found that although variance does not have a significant effect when mutation is neutral, the effect seems to be amplified when the mutation is slightly beneficial. In addition, we found that variance has different effects on structured networks that are not well-mixed. Specifically, we found that variance has a greater effect on star graphs as compared to well-mixed population. This project could be expanded in many different directions.

1. I would try more parameters for selective advantage s. This is to check if the amplification factor of network applies in general. I would also try some negative s corresponding to deleterious mutations, since it could be interesting, and most of the mutational that naturally occur are neutral to deleterious.

2. I would try running the simulation on more types of graphs including regular graphs and other graphs with special properties. This is to both confirm our current finding on network being properties being preserved in presence of variance, and to explore the effect on suppressor graphs.

3. Current simulation of variance are mostly using uniform and Bernoulli distributions for the ease of analytical analysis and ease of controlling variance. However, a bounded normal distribution would be more realistic according to figure 4.

4. So far, the fitness sampled during each birth-death process is independent and identically distributed. We would also like to explore the case when such fitness could be temporally or spatially correlated.

References

[1] Allen, P. M. Evolution, population dynamics, and stability. *Proceedings of the National Academy of Sciences 73*, 3 (1976), 665–668.

[2] Carja, O., and Feldman, M. W. An equilibrium for phenotypic variance in fluctuating environments owing to epigenetics. *Journal of The Royal Society Interface 9*, 69 (2012), 613–623.

[3] Carja, O., and Kuo, Y. Evolutionary graph theory beyond pairwise interactions: higher-order network motifs shape times to fixation in structured populations.

[4] Carja, O., Liberman, U., and Feldman, M. W. Evolution with stochastic fitnesses: A role for recombination. *Theoretical Population Biology 86* (2013), 29–42.

[5] Kuo, Y. P., Nombela-Arrieta, C., and Carja, O. A theory of evolutionary dynamics on any complex spatial structure. *bioRxiv* (2021).

[6] McAvoy, A., and Allen, B. Fixation probabilities in evolutionary dynamics under weak selection. *Journal of Mathematical Biology 82*, 3 (2021), 1–41.