

PH4209 Assignment — 02

22MS037

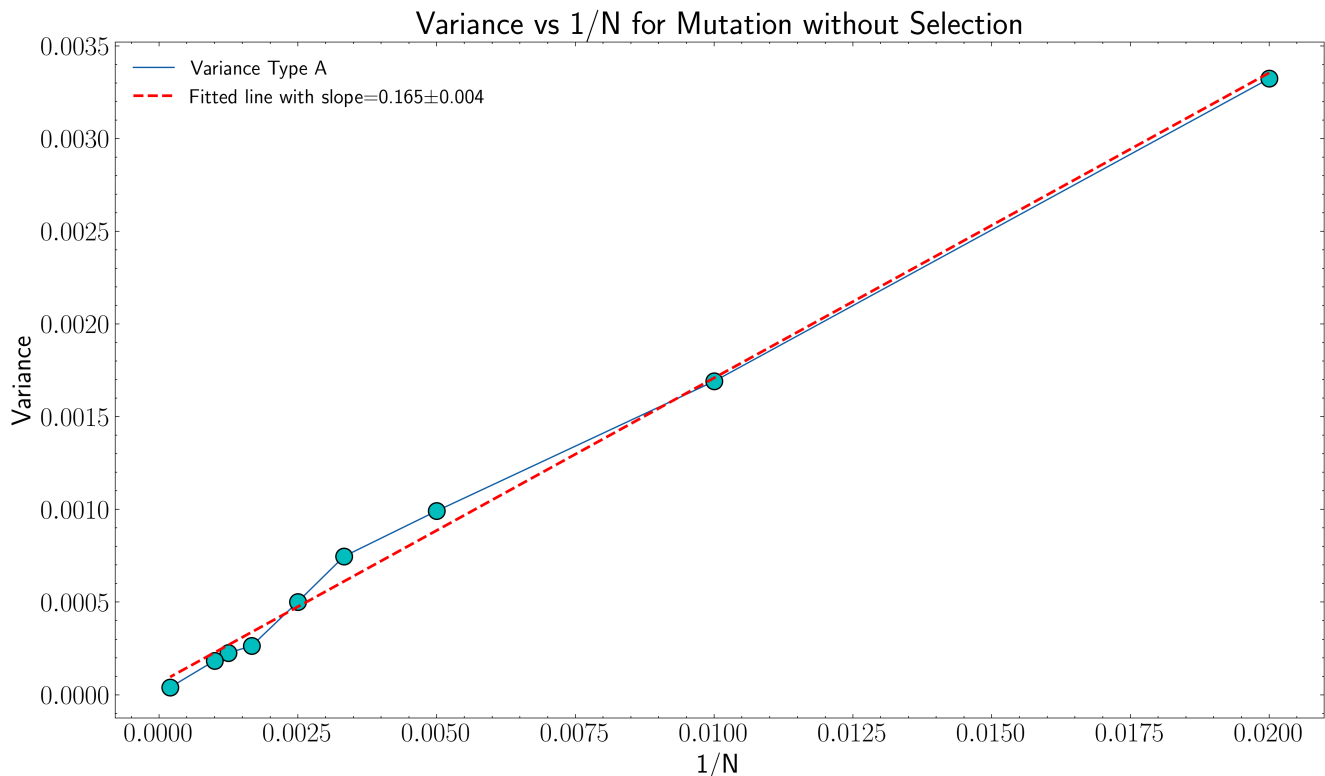
Problem 1 — Repeat Assignment 1 (mutation without selection) in a slightly different way. Divide population of size N equally between two types of individuals A and B represented by the numbers 0 & 1. Define mutation rates u_1 and u_2 . Use $u_1 = 0.003$ and $u_2 = 0.001$.

1. Repeat each simulation for different population sizes $N = 50, 100, 200, 300, 400, 600, 800, 1000, 5000$ for $N_T = 100$ trials
2. For each value of N , choose a time-step at which the system has equilibrated (use your results from Assignment 1 to make this choice). For that time step, calculate the mean frequency and variance of the frequency by averaging over $N_T = 100$ trials. Use the type with the large equilibrated frequency.

$$\langle f \rangle_{\text{trials}} = \frac{1}{N_T} \sum_{i=1}^{N_T} f \quad : \text{Mean} \quad (1)$$

$$(\delta f)_{\text{trials}}^2 = \langle f^2 \rangle_{\text{trials}} - (\langle f \rangle_{\text{trials}})^2 \quad : \text{Variance} \quad (2)$$

Plot a graph of Variance vs $(\frac{1}{N})$ and show that it is a straight line. This simulation is meant to show that fluctuations arising due to finite size of the population (quantified by variance) is inversely proportional to population size (N).



The time step chosen for equilibration is $T = 2000$, for all values of N . One can also code a dynamic check for equilibration based on when the variance of the frequency over the last few time steps becomes smaller than a threshold value. One assumes that if the larger population has equilibrated, the smaller ones would have equilibrated too.

The plot of Variance vs $(1/N)$ is shown above, along with the fitted straight line. The slope of the fitted line is 0.165 ± 0.004 , which shows that the variance is indeed inversely proportional to population size.

Problem 2 — Genetic Drift (neutral evolution) without mutation

Take a population of $N = 500$ individuals **half of which** consist of type 0 and the remaining half of type 1, initially. Assume that the fitness of both types are equal and neither type can mutate to the other.

1. Write a program to obtain the time-evolution of the frequencies of the two types in the population when individuals making up the population in the next generation are chosen randomly from members of the current generation following Moran process. Run the simulation for as long as it takes for any one of the two types to get fixed in the population.
2. Obtaining the Fixation Probability: Repeat the above simulation for $N_t = 100$ trials and find out the fraction of times each of the two types 0 and 1 get fixed? What do you expect the theoretical value of fixation probability for either sub-type to be?

For any one trial, plot the evolution of frequency of type 0 and type 1 with time

The Moran process simulates evolution one individual at a time. In our case, we are considering neutral evolution, where both types have equal fitness. Thus, at each time step, one individual is chosen randomly to reproduce, and one individual is chosen randomly to die, keeping the population size constant. **One edge case that is dealt with in the code is that if the same individual is chosen to reproduce and die, the population remains unchanged in that time step. This is sound modeling, as this scenario can be interpreted as different individuals of the same type chosen to die and to reproduce. However, this seems to be physically absurd. In our code, one individual can be chosen to both reproduce and die.**

The initial population is set up such that half the individuals are of type 0 and half are of type 1. The fixation probability for either type in neutral evolution is theoretically equal to its initial frequency in the population. Thus, for our case, the fixation probability for both types is expected to be 0.5.

$$\begin{aligned}x_A^* &= \frac{i}{N} = 0.5 \\x_B^* &= \frac{i}{N} = 0.5\end{aligned}\tag{3}$$

We run the simulation for $N_t = 100$ trials. The code is implemented using a for loop to ensure infinite loops are avoided. On an average for the given parameters, the average time to fixation is around 1.6×10^5 generations, averaged over 300 trials. So the maximum time steps is set to $T_{\max} = 10^6$ to be safe.

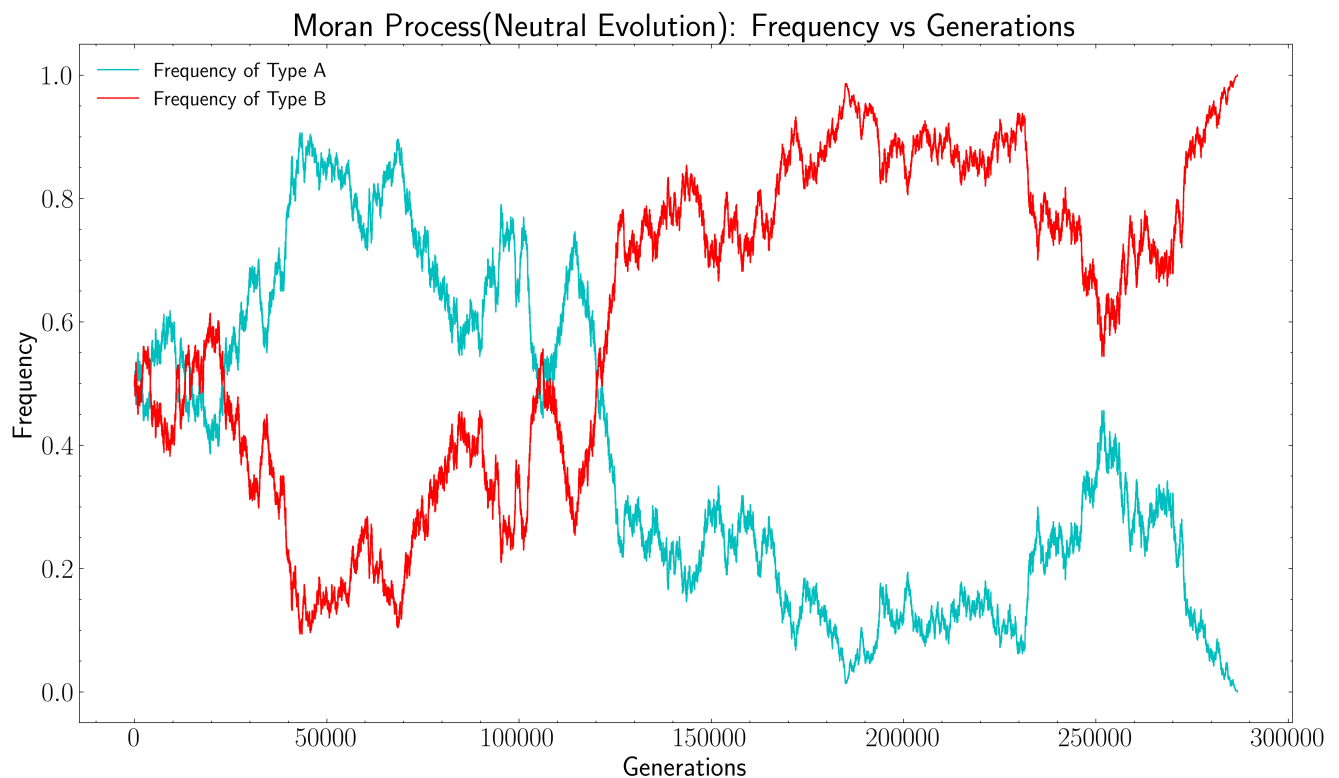


Figure 2: One realization for a Moran process for neutral evolution.

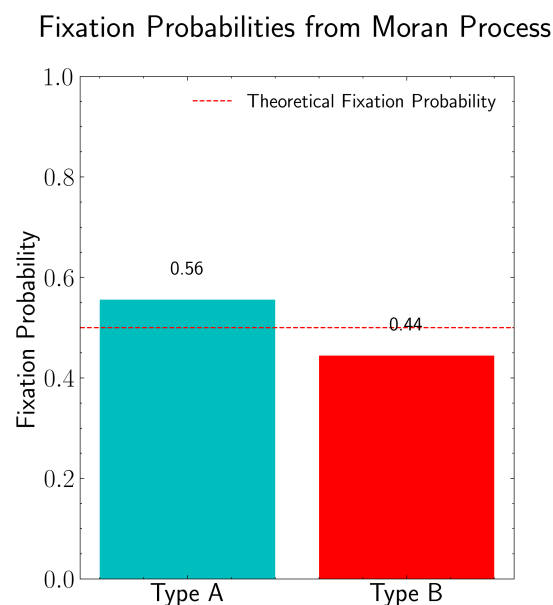


Figure 3: Fixation probability over 100 trials for Moran process for neutral evolution. In 56 out of 100 trials, type A got fixed, and in 44 out of 100 trials, type B got fixed. The theoretical fixation probability for both types is 0.5, which is close to the simulated result.