Evolutionary Population Dynamics

The goal of this lab is to explore evolutionary dynamics for a population with a fixed size and observe the effects of random genetic drift and selection. For this goal, we will employ variations of the Moran model (stochastic) and ODE population dynamics (deterministic) simulations. The motivation would be to consider the evolution of cancer in a small homeostatic tissue niche. Suppose a cancer cell invades (via metastasis) or stochastically arises (due to one or several mutations) is a certain tissue. If the resources or space are finite, cell growth in that tissue must be balanced by other cell death. What is the chance for the cancer cell to survive and perhaps eventually take over the population? The answer obviously depends on the population size and the relative growth advantage of the cancer cells. For your report, include all MATLAB code, figures generated, and answers to included questions.

Problem 1: Moran model with no selection (Demo)

The goal of this model is to investigate two algorithmic implementations of the Moran model, each of which would be useful for the subsequent parts.

General Moran model setup: Initially, we have N cells where $N-1$ are healthy and 1 is a mutant cancer cell. Cells divide asynchronously with generation time T_0, i.e. every time step $\Delta t = T_0/N$ there will be a cell randomly chosen to die and a second random cell is chosen to be replicate; in other words, the dead cell is replaced by the daughter of the dividing cell. Let the fitness of healthy and cancer cells be $g = \log(2)/T_0$ and $(1+s)g$, respectively, where s denotes the difference in fitness (i.e., $s > 0$ represents a selective advantage and $s < 0$ represents a selective disadvantage).

For Problem 1, we will simulate a situation in which there are no growth advantages for the cancer cells, i.e., $s = 0$. We will simulate the Moran model 1000 times for two different total populations $N = 10$ cells and $N = 100$ cells and examine the probability of cancer cells surviving after $T = 50$ days with a doubling time of $T_0 = 1\ day$. Your goal is to compare the simulations via the two algorithms below with one another and see how results change with N. For each algorithm A and B, you will plot the number of mutant cells M vs time (number of steps divided by Δt) for the first 3 simulations and the mean number of M for 1000 simulations. Plot the probability of survival for mutant cancer cells vs time and compute the probability of cancer cells surviving at the final time.

A) Algorithm A. Note that we can represent a cell population with a $1 \times N$ vector where healthy cells are denoted as 0 and cancer cells are denoted as 1. We simulate the model by changing $1 \rightarrow 0$ and $0 \rightarrow 1$ in the vector.
   1. Start by initializing a vector of N cells and randomly choose an initial index to place a single mutant cell.
   2. At each time step, generate a random integer index between 1 and N to select a cell to die and another random index to select a cell to reproduce and replace the first cell. Record the number of mutant cells (M) after each iteration in a vector.
   3. Repeat until the number of the time-steps is reached $T/\Delta t$ or all the mutants die ($M = 0$) or all the mutant cells take over the population ($M = N$).

B) In this algorithm we will not be tracking indices of the mutant cells but instead simply track their number. At each timestep, the number of mutant cells can change by 1 at most. For example, if there are M mutant cells at time t, we can have either $M-1$ or $M+1$ mutant cells at time $t + \Delta t$. The probability that the system will have n' individuals of a certain type in the next generation can be denoted as $T(n'|n)$. For a model where all individuals have the same fitness (i.e., there is no selection), we have

$$T(n'|n) = \begin{cases} \frac{n}{N}\left(1-\frac{n}{N}\right) & n' = n+1 \\ \frac{n}{N}\left(1-\frac{n}{N}\right) & n' = n-1 \\ 1 - T(n-1|n) - T(n+1|n) & n' = n \\ 0 & otherwise \end{cases}$$

Therefore, to simulate using the transition matrix approach,

1. Initialize a vector to store the number of mutants after each time step. Initially, we start with 1 mutant cell.
2. Generate a random number and use the probabilities from the transition matrix to determine if the number of mutants increases or decreases by 1.
   Repeat until the number of the time-steps is reached $T/\Delta t$ or all the mutants die ($M = 0$) or all the mutant cells take over the population ($M = N$).

Simulate 1000 realizations of the Moran model using the transition matrix and calculate the fraction of time a mutant cell still exists at the final time. Plot the number of mutant cells vs time for the first 3 simulations and the mean number of mutants vs time. Plot the probability of survival for mutant cancer cells vs time and calculate the probability of cancer cells surviving at the final time. Use $N = 10$ and 100 cells, $T_0 = 1$ day, $\Delta t = T_0/N$, and $T = 50$ days.

Problem 2: Moran model with selection

We will now build off problem 1 by giving mutant cancer cells a selective advantage by setting $s = 0.05$. We will employ the algorithm from 1B with the rescaled transition matrix for the Moran model with selection is

$$T(n'|n) = \begin{cases} (1+s)\frac{n}{N}\left(1-\frac{n}{N}\right) & n' = n+1 \\ \frac{n}{N}\left(1-\frac{n}{N}\right) & n' = n-1 \\ 1 - T(n-1|n) - T(n+1|n) & n' = n \\ 0 & otherwise \end{cases}$$

A) Simulate 1000 realizations of the number of mutants as a function of time using the rescaled transition matrix for the Moran model with $N = 10$ cells, $T_0 = 1$ day, $\Delta t = T_0/N$, $T = 150$ days, and $s = 0.05$. Plot the number of mutants vs time for the first 3 simulations and the mean number of mutants vs time. Plot the probability of survival for mutant cancer cells vs time and calculate the probability of cancer cells surviving at the final time.

B) Simulate the number of mutants as a function of time using the ODE

$$\frac{dM}{dt} = gM(1+s) - g_{av}M$$
$$g_{av} = \frac{M}{N}(1+s)g + \left(\frac{N-M}{N}\right)g$$

Use parameters $N = 10$ cells, $g = log(2)/T_0$, $s = 0.05$, $T_0 = 1$ day, $M_0 = 1$ and $T = 150$ days. Plot the solution of dM/dt vs time. How do these results compare to the stochastic simulations? Hint: what is the average number of mutants after 150 days? What is the final number of predicted mutants by the ODE?

C) Repeat part A-B) with $N = 100$ and compare your results to those with smaller N. Does increasing the number of cells significantly alter results? If so, how?

Problem 3: Spatial Moran model

A) We will now modify our model for the spatial effect. We can account for the spatial effect by
   1. Initialize a 1D lattice of N cells. Randomly choose a position on the lattice to place a mutant cancer cell.
   2. At each time step, randomly select a cell to die (index i).
   3. If the chosen cell (index i) is not on the boundary ($i = 1\ or\ N$) one of its neighbors ($i + 1$ or $i - 1$) is chosen to divide with probability proportional to their fitness. If both neighbors have the same fitness you select one of them with probability 0.5. Otherwise, a mutant is selected with probability $(1 + s)/(2 + s)$ and a wild-type cell is selected with probability $1/(2 + s)$.
   4. If the initially chosen cell dies on the boundary ($i = 1$ or $i = N$), the adjacent cell (respectively $i = 2$ or i=$N - 1$) divides. For example, if the cell is at $i = 1$ and dies, the cell at $i = 2$ divides.

   Simulate 1000 realizations of the spatial Moran model using $N = 10$ cells, $T_0 = 1$ day, $\Delta t = T_0/N$, $T = 50$ days and $s = 0.05$. Plot the number of mutants vs time for the first 3 simulations and the mean number of mutants vs time. Plot the probability of survival for mutant cancer cells vs time and calculate the probability of cancer cells surviving at the final time.
B) Repeat part A) with $N = 100$ and compare your results to those with smaller N.
C) Comment how mutant survival probability changes in this setup in comparison to spatially homogeneous situation in problem 2.