Anushka Sinha

02712 HW5

11/21/23

1. a.

$$\text{Minimize: } \sum_{i=1}^{N} \left(n_i - (u_i \times g) \right)^2$$

where N is total number of sites in the bacterial genome

n_i is the observed number of mutations at site i in the genome

u_i is the mutation rate at site i in the genome

g is the current generation

b. Probability of observing n mutations given a rate $\lambda = u_i g$ can be represented by the Poisson pdf:

$$P(n, u_i \times g) = \frac{e^{-(u_i \times g)} \times (u_i \times g)^n}{n!}$$

MLE formula:

$$\prod_{i=1}^{N} \frac{(u_i \times g)^{n_i}}{n_i\,!} e^{-(u_i \times g)}$$

where N is total number of sites in the bacterial genome

n_i is the observed number of mutations at site i in the genome

u_i is the mutation rate at site i in the genome

g is the current generation

② Bad State : 1

Good State : 2

Transition matrix = $P = \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix}$

p : probability of transitioning to good state

$1-p$: probability of remaining in the bad state

To find the stationary distribution π, solve $\pi P = \pi$

where π is a row vector = $[\pi_1, \pi_2]$

$$[\pi_1, \pi_2] \begin{bmatrix} 1-p & p \\ 1-p & p \end{bmatrix} = [\pi_1, \pi_2]$$

$$\pi_1(1-p) + \pi_2(1-p) = \pi_2$$

$$(\pi_1 + \pi_2)(1-p) = \pi_2$$

$$1(1-p) = \pi_2$$

$$1-p = \pi_2$$

$$\pi_1(p) + \pi_2(p) = \pi_1$$

$$(\pi_1 + \pi_2)p = \pi_1$$

$$1p = \pi_1$$

$$p = \pi_1$$

$\therefore$ Stationary distribution is $\pi = [p, 1-p]$

In this question, since the transition probability to the good state is independent of whether the system is currently in a good or bad state, one probability can describe the behavior of the system. Therefore, p becomes the only parameter that fully characterizes the system, and p alone can be used to express all stationary probabilities.

③ ⓐ $\det(A-\lambda I) =$

$$\begin{bmatrix} 1 & e_2 & 0 & 0 \\ 0 & (1-e_2)(1-b_2) & 0 & 0 \\ 0 & (1-e_2)b_2 & 1-e_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1-1 & e_2 & 0 & 0 \\ 0 & \lambda_2-[(1-e_2)(1-b_2)] & 0 & 0 \\ 0 & (1-e_2)b_2 & \lambda_3-(1-e_3) & 0 \\ 0 & 0 & 0 & \lambda_4-1 \end{bmatrix}$$

$$= (-1)^2(\lambda-1)\begin{vmatrix} \lambda_2-[(1-e_2)(1-b_2)] & 0 & 0 \\ (1-e_2)b_2 & \lambda_3-(1-e_3) & 0 \\ 0 & e_3 & \lambda_4-1 \end{vmatrix} + (-1)^3(e_2)\begin{vmatrix} 0 & 0 & 0 \\ 0 & \lambda_3-(1-e_3) & 0 \\ 0 & e_3 & \lambda_4-1 \end{vmatrix}$$

$$= (\lambda-1)\left((1)(\lambda-[(1-e_2)(1-b_2)]\begin{vmatrix} \lambda_3-(1-e_3) & 0 \\ e_3 & \lambda_4-1 \end{vmatrix}\right) + (e_2)\left((-1)^2(0)\begin{vmatrix} \lambda_3-(1-e_3) & 0 \\ e_3 & \lambda_4-1 \end{vmatrix}\right)$$

$$= (\lambda_1-1)\left[(\lambda_2-[(1-e_2)(1-b_2)])(\lambda_3-(1-e_3))(\lambda_4-1)\right] = 0$$

$\therefore \quad \lambda_1 - 1 = 0$

$\boxed{\lambda_1 = 1}$

$\lambda_2 - [(1-e_2)(1-b_2)] = 0$

$\boxed{\lambda_2 = (1-e_2)(1-b_2)}$

$\lambda_3 - (1-e_3) = 0$

$\boxed{\lambda_3 = 1-e_3}$

$\lambda_4 - 1 = 0$

$\boxed{\lambda_4 = 1}$

(b) The absorbing states are state 1 and state 4

(c) Transposed transition probability matrix

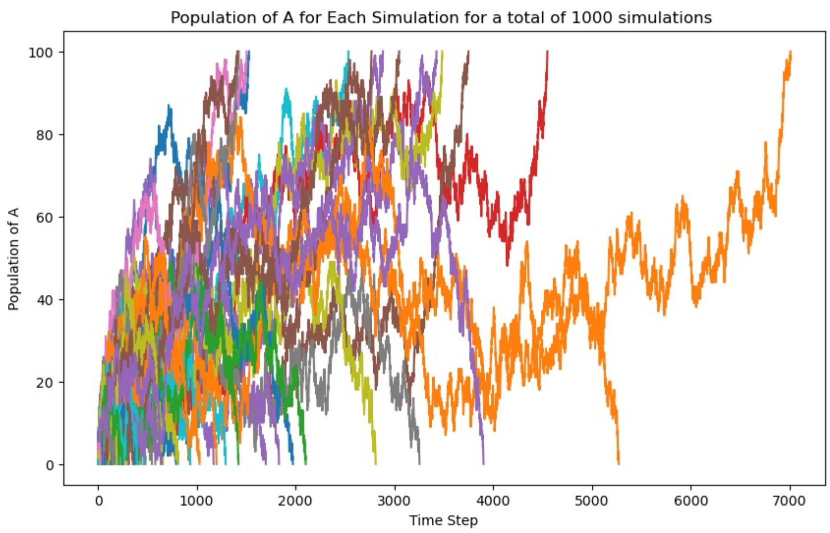
	(1)	(2)	(3)	(4)
(1)	1	0	0	0
(2)	e_2	$(1-e_2)(1-b_2)$	$(1-e_2)b_2$	0
(3)	0	0	$1-e_3$	e_3
(4)	0	0	0	1

Probability that the patch ultimately becomes a corn field (state 1) starting from state 2: $e_2 + e_2 \sum_{n=1}^{\infty} \left((1-e_2)(1-b_2)\right)^n$

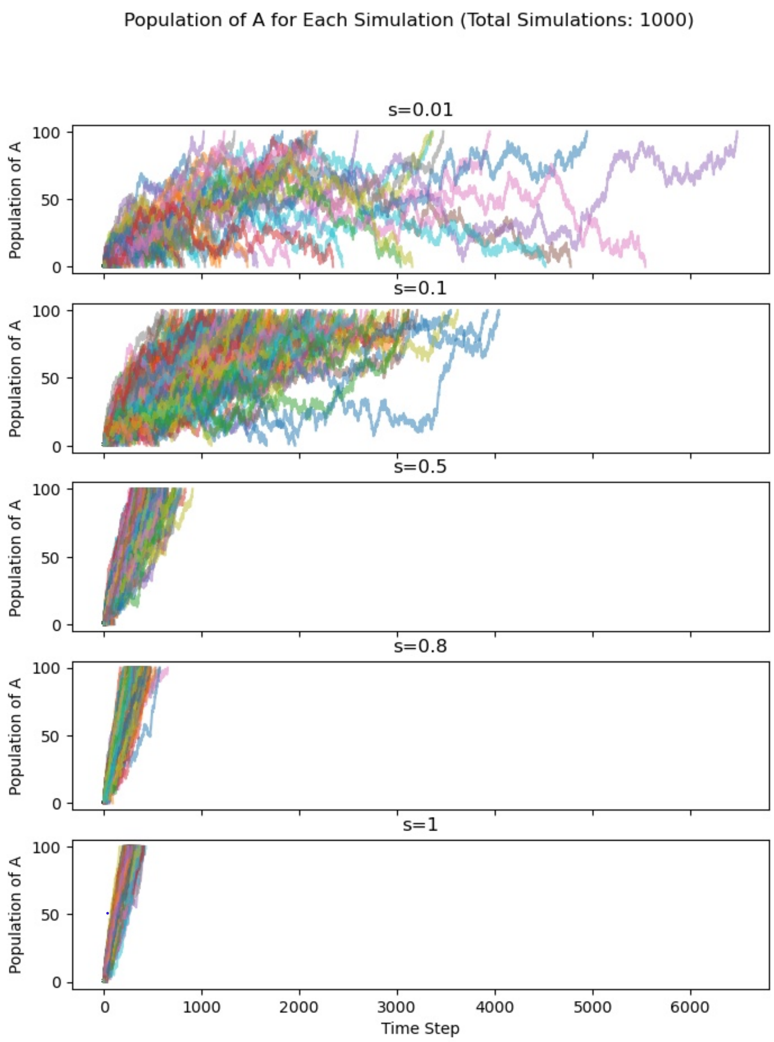
Probability that the patch ultimately becomes a corn field (state 1) starting from state 3: 0

Probability that the patch ultimately becomes a corn field (state 1) starting from state 2 or state 3: $e_2 + e_2 \sum_{n=1}^{\infty} \left((1-e_2)(1-b_2)\right)^n$

(4) (a)



b.



The plot above shows that allele A reaches the fixation point more frequently and quickly as the value of s (selective advantage) increases. This is indicated by the trajectories becoming steeper and getting more clumped together at higher values of A. Such an observation is consistent with the expectation that a larger selective advantage towards allele A should increase the probability of fixation of allele A.

5. (a) Transition Probabilities:

P(number of infections increasing by 1) = $\beta \delta I$

P(number of infections decreasing by 1) = $(c+d)I$

P(number of infections remaining the same) = $1-(\beta \delta I + (c+d)I)$

(b) $x_i = (c+d)I x_{i-1} + \beta\delta I x_{i+1} + (1-(\beta\delta I + (c+d)I)x_i$ $\quad (x_N = 1 \quad x_0 = 0)$

$x_i - (1-(\beta\delta I + (c+d)I)x_i = (c+d)I x_{i-1} + \beta\delta I x_{i+1}$

$$x_i = \frac{(c+d)I x_{i-1} + \beta\delta I x_{i+1}}{\beta\delta I + (c+d)I}$$

if $i = 1$ $\quad x_1 = \dfrac{\beta\delta I x_2}{\beta\delta I + (c+d)I}$

$\beta\delta I x_1 + (c+d)I x_1 = \beta\delta I x_2$

$(c+d)x_1 = \beta\delta(x_2 - x_1)$

when $\beta\delta > c+d$

$(x_2 - x_1)\beta\delta > (x_2 - x_1)(c+d)$

$(x_2 - x_1)\beta\delta > x_2(c+d) - x_1(c+d)$

$x_1(c+d) > x_2(c+d) - x_1(c+d)$

$2x_1(c+d) > x_2(c+d)$

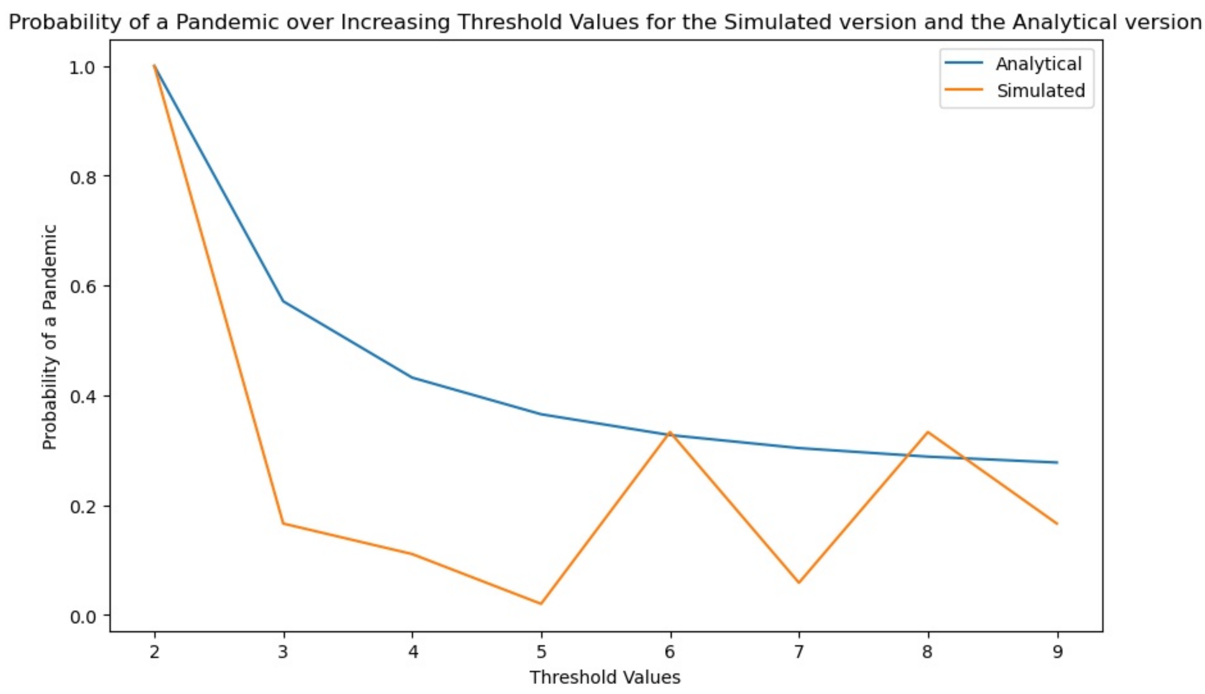
$2x_1 > x_2$

$2x_{n-1} > x_n$

when $x_n = 1$

$2x_{n-1} > 1$

$x_{n-1} > 0.5 \quad \therefore \quad 2x_{n-2} > x_{n-1} > 0.5$



The plot above shows that the simulated model displays stochasticity in the trajectory whereas the analytical solution maintains a smoother curve as it does not take into account any randomness.