Homework 5

Thomas Zhang

November 2023

Question 1

a)

$$\underset{u_i}{\text{argmin}} \sum_{i=1}^{m} (u_i g - n_i)$$

b)

$$\mathrm{P}(n_i|u_i, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5(\frac{n_i - \mu_i g}{\sigma})^2}$$

$$\mathcal{L}(u_{1:m}) = \prod_{i=1}^{m} \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5(\frac{n_i - \mu_i g}{\sigma})^2}$$

So the objective is:

$$\begin{aligned} &\underset{u_{1:m}}{argmax} \quad L(u_{1:m}) \\ =&\underset{u_{1:m}}{argmax} \quad \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5(\frac{n_i - \mu_i g}{\sigma})^2} \end{aligned}$$

Question 2

The stationary distribution of the two state system should be:

Good	Bad
p	$1-p$

$$\begin{bmatrix} p & 1-p \\ p & 1-p \end{bmatrix}^T \pi = \pi$$

$$\begin{bmatrix} p & p \\ 1-p & 1-p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{aligned} px_1 + px_2 &= x_1 \\ (1-p)x_1 + (1-p)x_2 &= x_2 \end{aligned}$$

$$\begin{aligned} (p-1)x_1 + px_2 &= 0 \\ x_1 &= \frac{px_2}{1-p} \end{aligned}$$

$$\begin{aligned} x_1 + x_2 &= 1 \\ \frac{px_2}{1-p} + x_2 &= 1 \\ px_2 + (1-p)x_2 &= (1-p) \\ x_2 &= 1-p \end{aligned}$$

Plugging in $x_2 = (1-p)$ we get $x_1 = p$.

The intuitive reason for this is that at each time point, the possibility of transitioning to the good state from any state is p. Thus after a long time interval, $p\%$ of the population should be in the good state.

Question 3

a)

$$\left\| \begin{bmatrix} 1 & e_2 & 0 & 0 \\ 0 & (1-e_2)(1-b_2) & 0 & 0 \\ 0 & (1-e_2)b_2 & 1-e_3 & 0 \\ 0 & 0 & e_3 & 1 \end{bmatrix} - \lambda I \right\| = 0$$

$$\left\| \begin{bmatrix} 1-\lambda & e_2 & 0 & 0 \\ 0 & (1-e_2)(1-b_2)-\lambda & 0 & 0 \\ 0 & (1-e_2)b_2 & 1-e_3-\lambda & 0 \\ 0 & 0 & e_3 & 1-\lambda \end{bmatrix} \right\| = 0$$

$$(1-\lambda)((1-e_2)(1-b_2)-\lambda)(1-e_3-\lambda)(1-\lambda)) = 0$$

Eigenvalues:

$$\lambda = 1$$
$$\lambda = 1$$
$$\lambda = (1-e_2)(1-b_2)$$
$$\lambda = 1-e_3$$

b)

State 1 and State 4 are the absorbing states because once a patch reaches state 1 or state 4 it does not leave.

c)

$$\mathrm{P}(S=1|S_0=3,t)=0$$

A patch starting from state 3 has a 0% chance to reach state 1.

$$\begin{aligned}\mathrm{P}(S=1|S_0=2) &= \sum_{i=1}^{\infty}((1-e_2)(1-b_2))^{i-1}e_2\\ &= e_2\sum_{i=1}^{\infty}((1-e_2)(1-b_2))^{i-1}\end{aligned}$$

Let $x=(1-e_2)(1-b_2)$:

$$\begin{aligned}\sum_{i=1}^{\infty}x^{i-1} &= 1+x+\dots\\ x\sum_{i=1}^{\infty}x^{i-1} &= x+x^2+\dots\\ x\sum_{i=1}^{\infty}x^{i-1}-\sum_{i=1}^{t}x^{i-1} &= -1\\ (x-1)\sum_{i=1}^{\infty}x^{i-1} &= -1\\ \sum_{i=1}^{\infty}x^{i-1} &= \frac{-1}{x-1}\\ \sum_{i=1}^{\infty}x^{i-1} &= \frac{1}{1-x}\end{aligned}$$

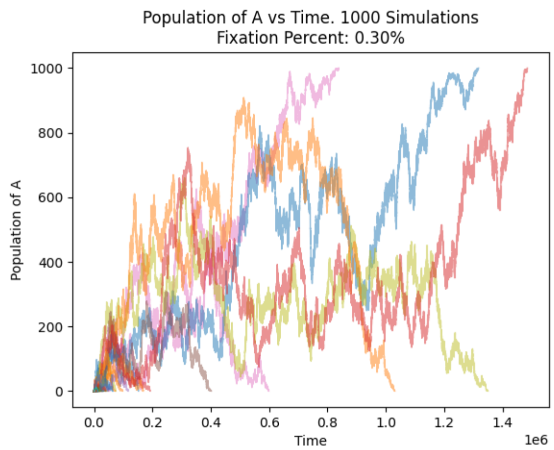
$$\begin{aligned}\mathrm{P}(S=1|S_0=2) &= e_2\sum_{i=1}^{\infty}((1-e_2)(1-b_2))^{i-1}\\ &= \frac{e_2}{1-(1-e_2)(1-b_2)}\end{aligned}$$

The probability of reaching state 1 of a patch starting at state 2 is:

$$\frac{e_2}{1-(1-e_2)(1-b_2)}$$

Question 4

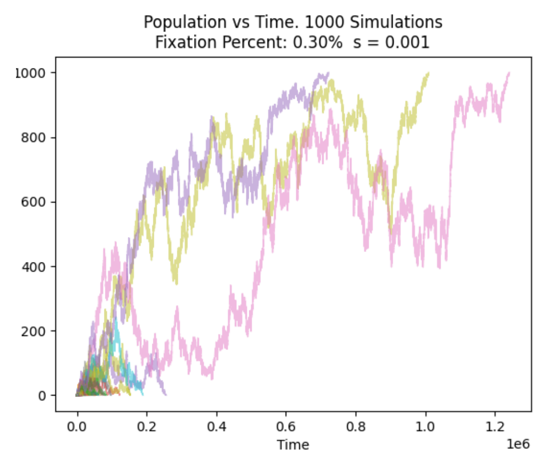
a)

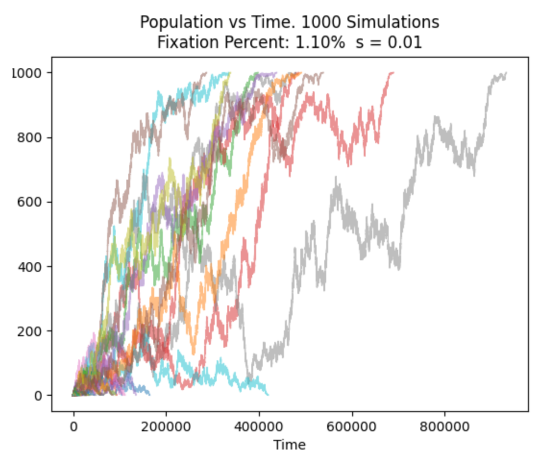


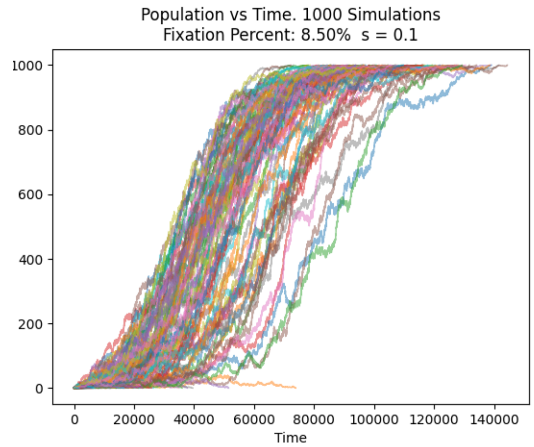
The simulation of the Moran process resulted in a fixation rate approximately equal to the theoretical fixation rate of i_0/N where i_0 is the number of gene A at time 0. In this case $i_0 = 1$ and $N = 1000$ so the theoretical fixation rate is $0.001 = 0.1\%$. 1000 simulations returned a fixation rate of 0.3% which is close to the theoretical.

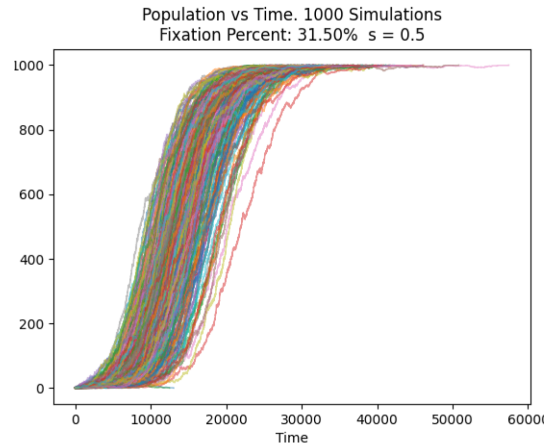
b)

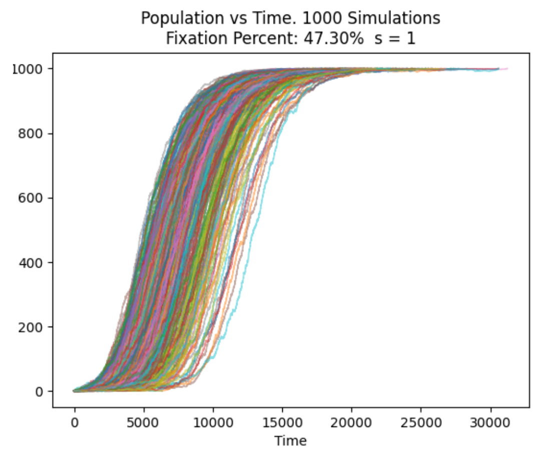
Simulations with $A_0 = 1$.







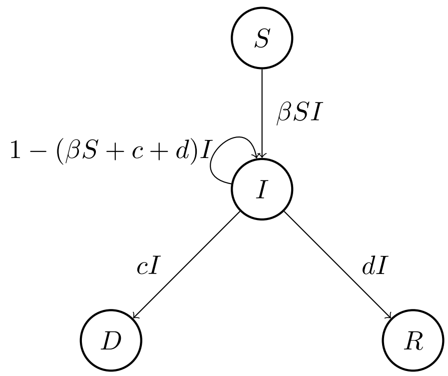




Based on the plots, as the selective benefit, s, for A increases, the fixation rate of A increases. This makes sense because as s increases, the reproductive advantage of A increases. Hence, A will be favored to reach fixation.

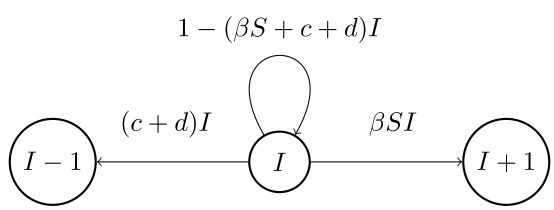
Question 5

a)



	S	I	D	R
S	$1 - \beta SI$	βSI	0	0
I	0	$1 - (\beta S + c + d)I$	cI	dI

Table 1: Transition Probabilities



	I-1	I	I+1
I	$(c+d)I$	$1 - (\beta S + c + d)I$	βSI

Table 2: Transition Probabilities

b)

$$x_0 = 0 \tag{1}$$
$$x_i = (c+d)ix_{i-1} + (1 - (\beta S + c + d)i)x_i + \beta Six_{i+1} \tag{2}$$
$$x_m = 1 \tag{3}$$

Evaluating equation 2.

$$\begin{aligned} x_i &= (c+d)ix_{i-1} + (1-(\beta S + c + d)i)x_i + (\beta S)ix_{i+1} \\ 0 &= (c+d)ix_{i-1} - (\beta S + c + d)ix_i + (\beta S + c + d)ix_{i+1} \\ 0 &= (c+d)x_{i-1} - (\beta S + c + d)x_i + (\beta S)x_{i+1} \end{aligned}$$

Let $u = c + d$ and $v = \beta S$.

$$\begin{aligned} 0 &= ux_{i-1} - (u+v)x_i + vx_{i+1} \\ 0 &= ux_{i-1} - ux_i - vx_i + vx_{i+1} \\ ux_i - ux_{i-1} &= vx_{i+1} - vx_i \\ u(x_i - x_{i-1}) &= v(x_{i+1} - x_i) \\ \frac{u}{v}(x_i - x_{i-1}) &= (x_{i+1} - x_i) \end{aligned}$$

Let $y_i = x_i - x_{i-1}$

$$\frac{u}{v} y_i = y_{i+1} \tag{4}$$

Using the definition of y_i we can derive:

$$\begin{aligned} \sum_{i=1}^{j} y_i &= x_j - x_{j-1} + x_{j-1} - x_{j-2} + ... + x_1 - x_0 \\ &= x_j - x_0 \\ &= x_j \end{aligned} \tag{5}$$

Next using (4) we can get an expression for y_k where $k \leq m$.

$$\begin{aligned} y_k &= y_{k-1}(\frac{u}{v}) \\ &= y_{k-2}(\frac{u}{v})(\frac{u}{v}) \\ &= \dots \\ &= y_1 \prod_{i=1}^{k-1} \frac{u}{v} \\ &= (x_1 - x_0)\left(\frac{u}{v}\right)^{k-1} \\ &= x_1 \left(\frac{u}{v}\right)^{k-1} \end{aligned} \tag{6}$$

Now using (6) we get evaluate $\sum_{i=1}^{j} y_i$ again to get an expression different from what we derived earlier in (5):

$$\begin{aligned}\sum_{i=1}^{j} y_i &= y_1 + y_2 + \ldots + y_j \\ &= x_1 + x_1\left(\frac{u}{v}\right) + x_1\left(\frac{u}{v}\right)^2 + \ldots + x_1\left(\frac{u}{v}\right)^{j-1} \\ &= x_1(1 + \left(\frac{u}{v}\right) + \left(\frac{u}{v}\right)^2 + \ldots + \left(\frac{u}{v}\right)^{j-1}) \\ &= x_1(1 + \sum_{i=1}^{j-1}\left(\frac{u}{v}\right)^i)\end{aligned} \tag{7}$$

Now combine equations (5) and (7):

$$\begin{aligned}\sum_{i=1}^{m} y_i &= x_m && \text{from eq. (5)} \\ \sum_{i=1}^{m} y_i &= x_1(1 + \sum_{i=1}^{m-1}\left(\frac{u}{v}\right)^i) && \text{from eq. (7)} \\ x_m &= x_1(1 + \sum_{i=1}^{m-1}\left(\frac{u}{v}\right)^i)\end{aligned}$$

From our initial condition (equation 3), we know $x_m = 1$.

$$\begin{aligned}1 &= x_1(1 + \sum_{i=1}^{m-1}\left(\frac{u}{v}\right)^i) \\ x_1 &= \frac{1}{1 + \sum_{i=1}^{m-1}\left(\frac{u}{v}\right)^i}\end{aligned}$$

Now we can solve for any x_j for $1 < j \leq m$ using equations (5), (7) and the expression for x_1 we just derived.

$$\begin{aligned}\sum_{i=1}^{j} &= x_j && \text{From eq. (5)} \\ \sum_{i=1}^{j} &= x_1(1 + \sum_{i=1}^{j-1}\left(\frac{u}{v}\right)^i) && \text{From eq. (7)} \\ x_j &= x_1(1 + \sum_{i=1}^{j-1}\left(\frac{u}{v}\right)^i) \\ x_j &= \frac{1 + \sum_{i=1}^{j-1}\left(\frac{u}{v}\right)^i}{1 + \sum_{i=1}^{m-1}\left(\frac{u}{v}\right)^i}\end{aligned}$$

Deriving closed form of $\sum_{i=1}^{n} a^i$.

$$\begin{aligned}
\sum_{i=1}^{n} a^i &= a + a^2 + \ldots + a^n \\
a \sum_{i=1}^{n} a^i &= a^2 + a^3 + \ldots + a^{n+1} \\
(a-1) \sum_{i=1}^{n} a^i &= a^{n+1} - a \\
\sum_{i=1}^{n} a^i &= \frac{a^{n+1} - a}{a-1}
\end{aligned}$$

So the probability of a pandemic given we start at I = 1 is:

$$\begin{aligned}
\mathrm{P}(\text{Pandemic} \mid I_0 = 1) = x_1 &= \frac{1}{1 + \sum_{i=1}^{m-1} \left(\frac{u}{v}\right)^i} \\
&= \frac{1}{1 + \frac{\left(\frac{u}{v}\right)^m - \left(\frac{u}{v}\right)}{\left(\frac{u}{v}\right) - 1}} \\
&= \frac{\left(\frac{u}{v}\right) - 1}{\left(\frac{u}{v}\right) - 1 + \left(\frac{u}{v}\right)^m - \left(\frac{u}{v}\right)} \\
&= \frac{\left(\frac{u}{v}\right) - 1}{\left(\frac{u}{v}\right)^m - 1}
\end{aligned}$$

Plugging u and v back in we get:

$$\begin{aligned}
\mathrm{P}(\text{Pandemic} \mid I_0 = 1) &= \frac{1}{1 + \sum_{i=1}^{m-1} \left(\frac{c+d}{\beta S}\right)^i} \\
&= \frac{\left(\frac{c+d}{\beta S}\right) - 1}{\left(\frac{c+d}{\beta S}\right)^m - 1} \\
&= \frac{1 - \left(\frac{c+d}{\beta S}\right)}{1 - \left(\frac{c+d}{\beta S}\right)^m}
\end{aligned}$$

From looking at the plots it is hard to discern the relationship between the fixation rates and the parameters. It appears that the fixation rate is proportional to βS and inversely proportional to $c + d$.

One strange observation from the simulation is that the fixation rate is not really affected by the pandemic threshold. However looking at the closed form of the fixation rate, this is most likely because when m is large, the denominator term in

$$\frac{1 - \left(\frac{c+d}{\beta S}\right)}{1 - \left(\frac{c+d}{\beta S}\right)^m}$$

evaluates to approximately 1 given that $c + d < \beta S$. The m-terms tested in the simulation were 100 to 1000, which means that the denominator is approximately 1 across all m's.