3

problem 5

- 2N genes (population of constant size N diploid organisms)
- at time $n=0$, x of these genes are type A $(0 \le x \le 2N)$
- $(X(t) \mid X(t-1)=i) \sim Bin(2N, i/2N)$
- $$P(X(t)=j \mid X(t-1)=i) = \frac{2N!}{j!\,(2N-j)!} \left(\frac{i}{N}\right)^{j} \left(1-\frac{i}{N}\right)^{N-j}$$

a we know that:

$$E[y_n] = E[E(y_n \mid y_{n-1})] = E(y_{n-1}) = E(y_{n-2}) = \cdots = E(y_0) \quad (1)$$

$\Rightarrow$ $E[X(t) \mid X(0)=i] = Np = N(i/N) = i$ and by doing (1) repeatedly

$E[X(t) \mid X(t-1)=i] = i$

b

- if $X \sim Bin(N,p)$ then $Var(X) = Np(1-p)$
- also we know $Var(Y) = E[Var(Y|X)] + Var[E(Y|X)] \quad (2)$

$$\begin{aligned}
Var(X_t) &= E[Var(X_t|X_{t-1})] + Var[E(X_t|X_{t-1})] \\
&= E\left[2N \frac{X_{t-1}}{2N}\left(1-\frac{X_{t-1}}{2N}\right)\right] + Var(X_{t-1}) \\
&= E(X_{t-1}) - \frac{1}{2N} E(X_{t-1}^2) + Var(X_{t-1}) \\
&= E(X_{t-1}) - \frac{1}{2N}\left[Var(X_{t-1}) + E(X_{t-1})^2\right] + Var(X_{t-1}) \quad (3)
\end{aligned}$$

by assuming $X_t = (X(t)|X(0)=i)$ and $X_{t-1} = (X(t-1)|X(0)=i)$ we have

$$\begin{aligned}
Var[X(t)|X(0)=i] &= E[X(t-1)|X(0)=i] - \frac{1}{2N}\left[Var(X(t-1)|X(0)=i) + E[X(t-1)|X(0)=i]^2\right] + \\
&\quad Var[X(t-1)|X(0)=i] \\
&= i - \frac{1}{2N} Var(X(t-1)|X(0)=i) - \frac{1}{2N} i^2 + Var(X(t-1)|X(0)=i) \quad (4)
\end{aligned}$$

$$\Rightarrow Var[X(t)|X(0)=i] = \left(1-\frac{1}{2N}\right) Var(X(t-1)|X(0)=i) + i\left(1-\frac{i}{2N}\right) \quad (5)$$

if we assume that total size of population is N (instead of 2N)

then $\Rightarrow Var[X(t) \mid X(0)=i] = (1-\frac{1}{N})\, Var(X(t-1) \mid X(0)=i) + \underbrace{i(1-\frac{i}{N})}_{V_1}$ (6)

therefore, (from 4)

$$Var(X_t) - 2Ni(1-\frac{i}{2N}) = (1-\frac{1}{2N})\, Var(X_{t-1}) + i(1-\frac{i}{2N}) - 2Ni(1-\frac{i}{2N})$$

$$= (1-\frac{1}{2N})\, Var(X_{t-1}) + i - \frac{i^2}{2N} - 2Ni + i^2$$

$$= (1-\frac{1}{2N})\left[Var(X_{t-1}) - 2Ni(1-\frac{i}{2N})\right] \quad (7)$$

with boundary condition $X(0)=i$ we finally have:

$$Var(X_t) = 2Ni(1-\frac{i}{2N})\left[1-(1-\frac{1}{2N})^t\right] \quad (8)$$

again, if we assume that total population size is N

we have: $Var(X_t) = Var[X(t) \mid X(0)=i]$

$$Var(X_t) = Ni(1-\frac{i}{N})\left[1-(1-\frac{1}{N})^t\right] \quad (9)$$

$\boxed{C}$ $N \to \infty$ $\Rightarrow \lim_{N\to\infty} Var(X_t) = \lim_{N\to\infty} Ni(1)(1-1) = 0$

Since we have full matrix (instead of tri diagonal matrix in Moran)

also we look at n generation instead of only one generation.