Stochastic Processes

Assignment #2, Solutions

Total Marks: 39 for 455 and 45 for 855.

1. From Sheet. (6 marks) In the following solutions we use the result derived in class that for a Galton-Watson branching process the probability of ultimate extinction is given by the smallest nonnegative solution of the equation \( s = G(s) \), where \( G(\cdot) \) is the probability generating function of the family size distribution.

(a) (3 marks) For the Geometric(\( p \)) distribution as given in the problem, the generating function of the family size distribution is computed as

\[
G(s) = \sum_{n=0}^{\infty} s^n p q^n = p \sum_{n=0}^{\infty} (qs)^n = \frac{p}{1 - qs},
\]

which is valid for \( s \neq 1/q \), where \( q = 1 - p \). Therefore, the equation \( s = G(s) \) reduces to \( qs^2 - s + p = 0 \), which has solutions

\[
s = \frac{1 \pm \sqrt{1 - 4pq}}{2q} = \frac{1 \pm |1 - 2q|}{2q} = \frac{1 \pm (2q - 1)}{2q},
\]

since \( q > 1/2 \). Therefore, the solutions are \( s = 1 \) and \( s = p/q < 1 \), and the smallest solution, which is the probability of ultimate extinction, is \( p/q \).

(b) (3 marks) For this part the generating function of the family size distribution is given by

\[
G(s) = s^0 \left( \frac{1}{4} \right) + s^1 \left( \frac{1}{4} \right) + s^2 \left( \frac{1}{2} \right),
\]

and so the equation \( s = G(s) \) reduces to \( 2s^2 - 3s + 1 = 0 \), which has solutions

\[
s = \frac{3 \pm \sqrt{9 - 8}}{4} = \frac{3 \pm 1}{4}.
\]

Therefore, the solutions are \( s = 1 \) and \( s = 1/2 \), and the probability of ultimate extinction is \( 1/2 \).
2. **From Sheet.** (6 marks)

(a) (2 marks) States 1, 2, 5 and 6 are recurrent; states 3 and 4 are transient. All states have period 1 since $p_{ii}(1) > 0$ for all $i$. The equivalence classes are $\{1, 2\}$, $\{3, 4\}$, and $\{5, 6\}$.

(b) (4 marks) The recurrent states are 1, 2, 5 and 6, and $\{1, 2\}$ and $\{5, 6\}$ are closed recurrent classes. For state 1 the only way for the first return to be at time $n$ is to be in state 2 at times $1, \ldots, n - 1$ and then to go back to state 1 at time $n$. This gives us

$$f_{11}(n) = p_{12}p_{22}^{n-2}p_{21}$$

for $n \geq 2$, and $f_{11}(1) = p_{11}$. A similar analysis holds for states 2, 5 and 6. This gives us

$$f_{11}(n) = p_{12}p_{22}^{n-2}p_{21} = \left(\frac{1}{2}\right)\left(\frac{3}{4}\right)^{n-2}\left(\frac{1}{4}\right)$$

for $n \geq 2$, and $f_{11}(1) = p_{11} = \frac{1}{2}$

$$f_{22}(n) = p_{21}p_{11}^{n-2}p_{12} = \left(\frac{1}{4}\right)^{n-2}\left(\frac{1}{2}\right)$$

for $n \geq 2$, and $f_{22}(1) = p_{22} = \frac{3}{4}$

$$f_{55}(n) = p_{56}p_{66}^{n-2}p_{65} = \left(\frac{1}{2}\right)^{n-2}\left(\frac{1}{2}\right)$$

for $n \geq 2$, and $f_{55}(1) = p_{55} = \frac{1}{2}$

$$f_{66}(n) = p_{65}p_{55}^{n-2}p_{56} = \left(\frac{1}{2}\right)^{n-2}\left(\frac{1}{2}\right)$$

for $n \geq 2$, and $f_{66}(1) = p_{66} = \frac{1}{2}$

The mean return time to state $i$ is $\mu_i = \sum_{n=1}^{\infty} f_{ii}(n)$. From the above we have

$$\mu_1 = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} (n+1) \left(\frac{1}{4}\right)^n = \frac{1}{2} + \frac{1/2}{1/4} + \frac{1/8}{1 - 3/4} = 3$$

$$\mu_2 = \frac{3}{4} + \frac{1}{4} \sum_{n=1}^{\infty} (n+1) \left(\frac{1}{2}\right)^n = \frac{3}{4} + \frac{1/2}{1/2} + \frac{1/8}{1 - 1/2} = \frac{5}{2}$$

$$\mu_5 = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} (n+1) \left(\frac{1}{2}\right)^n = \frac{1}{2} + \frac{1/2}{1/2} + \frac{1/4}{1 - 1/2} = 2$$

$$\mu_6 = \frac{1}{2} + \frac{1}{2} \sum_{n=1}^{\infty} (n+1) \left(\frac{1}{2}\right)^n = \frac{1}{2} + \frac{1/2}{1/2} + \frac{1/4}{1 - 1/2} = 2$$

3. **From Sheet.** (9 marks) We first consider the symmetric simple random walk in two dimensions, i.e., $p_1 = p_2 = p_3 = p_4 = 1/4$. As suggested in the hint we compute $P(X_{2n} = (0, 0) \mid X_0 = (0, 0))$. In going from state $(0, 0)$ back to state $(0, 0)$ in $2n$ moves, the walk must make $2m$ moves in the vertical direction (either up or down),
for some $m \in \{0, 1, \ldots, n\}$, and the other $2(n - m)$ moves must be in the horizontal direction (either left or right). For a given $m$, it must further make $m$ moves up, $m$ moves down, $n - m$ moves left, and $n - m$ moves right, and there are $\binom{2n}{2m} \binom{2m}{m} \binom{2(n-m)}{n-m}$ distinct sample path segments that do this (equal to the number of ways to choose the $2m$ time points at which to move vertically times the number of ways to choose from these $2m$ time points the time points at which to move up times the number of ways to choose from the remaining $2(n - m)$ time points the time points at which the walk moved right). Each such sample path segment (still for a given $m$) has probability $p_m^2 p_{n-m}^2 = \left(\frac{1}{4}\right)^{2n}$. Altogether, we have

$$P(X_{2n} = (0, 0) \mid X_0 = (0, 0)) = \sum_{m=0}^{n} \binom{2n}{2m} \binom{2m}{m} \binom{2(n-m)}{n-m} \left(\frac{1}{4}\right)^{2n}$$

$$= \sum_{m=0}^{n} \frac{(2n)!}{m!m!(n-m)!(n-m)!} \left(\frac{1}{4}\right)^{2n}$$

$$= \sum_{m=0}^{n} \frac{(2n)!n!n!}{n!m!m!(n-m)!(n-m)!} \left(\frac{1}{4}\right)^{2n}$$

$$= \left(\frac{2n}{n}\right) \left(\frac{1}{4}\right)^{2n} \sum_{m=0}^{n} \binom{n}{m}^2$$

$$= \left(\frac{2n}{n}\right) \left(\frac{1}{4}\right)^{2n}$$

using the identity $\sum_{m=0}^{n} \binom{n}{m}^2 = \binom{2n}{n}$. Using Stirling’s approximation we have

$$P(X_{2n} = (0, 0) \mid X_0 = (0, 0)) \approx \left(\frac{(2n)^{2n+1/2}e^{-2n}\sqrt{2\pi}}{n^{n+1/2}e^{-n}\sqrt{2\pi n^{n+1/2}e^{-n}\sqrt{2\pi}4^n}}\right)^2$$

$$= \left(\frac{2^{2n+1/2}}{n^{1/2}\sqrt{2\pi}4^n}\right)^2$$

$$= \frac{1}{n\pi}. $$

Thus, since $\sum_{n=1}^{\infty} \frac{1}{n}$ is a divergent sum,

$$\sum_{n=0}^{\infty} p_{(0,0),(0,0)}(2n) = 1 + \sum_{n=1}^{\infty} P(X_{2n} = (0, 0) \mid X_0 = (0, 0))$$

$$\approx 1 + \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} = \infty,$$

and so state $(0, 0)$ (and, by irreducibility, every state) is recurrent by a proposition from class.
For the case $p_1 \neq p_3$, first map each state $(i, j)$ to $i + j$, the sum of its components. In other words, for each integer $k$, aggregate all states $(i, j)$ satisfying $i + j = k$ into a single state $k$. On this aggregated state space, from any state $k$ our random walk moves from state $k$ to state $k + 1$ with probability $p_u = p_1 + p_4$ and moves from state $k$ to state $k - 1$ with probability $p_d = p_2 + p_3$ (i.e., it is a one-dimensional simple random walk on the aggregated state space). If we show that $p_u \neq p_d$ then we know from class that this one-dimensional random walk is transient (i.e., all states $k$ are transient). Then this implies the original two-dimensional random walk is transient. This is so because starting in state $(0, 0)$ there is a positive probability of never returning to the line $x + y = 0$. But never returning to this line implies never returning to $(0, 0)$, so the probability of never returning to $(0, 0)$ is also positive, which makes state $(0, 0)$ transient, and by irreducibility all states are transient. It is possible that $p_u = p_1 + p_4 = 1/2$. However, if this is the case then it must be that $p_3 + p_4 \neq 1/2$, since $p_1 \neq p_3$. In this case we can apply the same argument as before except this time aggregating all states $(i, j)$ that satisfy $j - i = k$ into the single state $k$ (that is, consider the one-dimensional random walk on the diagonal lines with slope $+1$). State $0$ will be transient and so state $(0, 0)$ in the original two-dimensional random walk will be transient.

4. From Sheet. (6 marks) The state space is the set of nonnegative integers. Suppose the current polygon has $k$ sides, (so that the process is currently in state $k - 3$). When we choose two of these edges at random and join their midpoints we form two new polygons, one with $3 + E_1$ edges and one with $3 + E_2$ edges, where the 3 edges come from the two edges we chose plus the one edge joining their midpoints, and $E_1$ and $E_2$ are the numbers of edges between the chosen edges, counting in the two possible directions, respectively. Note that $E_1 + E_2 = k - 2$ and the possible values of $E_1$ are $0, 1, \ldots , k - 2$. Suppose we want the next polygon in the sequence to have $m$ edges, where the possible values of $m$ are $3, \ldots , k + 1$. For this to occur we need at least one of $E_1$ or $E_2$ to equal $m - 3$. If $m - 3 = (k - 2)/2$ (this corresponds to $E_1 = E_2$ and can only occur if $k$ is even) then there are $k/2$ distinct (unordered) edge pairs that give $E_1 = E_2 = (k - 2)/2$. In this case both polygons formed by choosing one of these edge pairs will have $m$ edges, and we get

$$P(X_{n+1} = m - 3 \mid X_n = k - 3) = P(X_{n+1} = k/2 - 1 \mid X_n = k - 3) = \frac{k/2 \choose (k/2)}{k} = \frac{1}{k - 1}.$$ 

For every other possible value of $m$, if one of the polygons has $m$ sides then the other will have $k + 4 - m \neq m$ sides. In this case there are $k$ distinct (unordered) edge pairs that give rise to one of the polygons having $m$ edges, and with probability $1/2$ we will
choose that polygon as the next one in the sequence. Therefore, for $m - 3 \neq (k - 2)/2$, we have

$$P(X_{n+1} = m - 3 \mid X_n = k - 3) = \frac{k}{(k/2)} \times \frac{1}{2} = \frac{1}{k - 1}.$$  

In other words, if the current polygon has $k$ sides, the next polygon in the sequence is equally likely to have $3, 4, \ldots, k + 1$ sides. The transition matrix of the Markov chain $\{X_n : n \geq 0\}$ is given by

$$P = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 & \cdots \\ 1/3 & 1/3 & 1/3 & 0 & 0 & \cdots \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

where the rows (and columns) correspond to the states $0, 1, 2, \ldots$.

5. From Sheet. (12 marks) We note here that $d$ distinct states, $i_0, \ldots, i_{d-1}$, as described in the problem statement, must exist because, if not, then in the first $d$ states visited by the chain at least one of these states must be visited at least twice, which contradicts the assumption that the period of all states is $d$.

(a) (3 marks) Assume that $i \in C_j$. We know that $p_{i,i_j}(m) > 0$ for some $m$ because $i$ communicates with $i_j$. Also, by definition of $C_j$, there is some $n_0$ such that $p_{i_j,i}(n_0d) > 0$. Therefore,

$$p_{ii}(m + n_0d) \geq p_{i,i_j}(m)p_{i_j,i}(n_0d) > 0.$$  

This implies $m$ must be of the form $nd$ for some $n$ since otherwise the fact that state $i$ has period $d$ is contradicted. Conversely, suppose that $p_{i,i_j}(n_0d) > 0$ for some $n_0$. Again, since $i$ communicates with $i_j$, there is some $m$ such that $p_{i_j,i}(m) > 0$ so that

$$p_{ii}(n_0d + m) \geq p_{i,i_j}(n_0d)p_{i_j,i}(m) > 0,$$

and the period of state $i$ being $d$ is again contradicted if $m$ is not of the form $nd$.

(b) (3 marks) Suppose there exists a state $i$ that belongs to both $C_j$ and $C_k$, where $j < k$. By construction, there is a path segment from state $i_j$ to state $i_k$ that has probability

$$p_{i_j,i_{j+1}} \cdots p_{i_{k-1},i_k} > 0$$
with length \( k - j \in \{1, \ldots, d - 1\} \). Since \( i \in C_j \), by part(a) there is some \( m \) such that \( p_{i,i_j}(md) > 0 \). Since \( i \in C_k \), by definition there is some \( n \) such that \( p_{i,k}(nd) > 0 \). But then

\[
p_{i}(md + k - j + nd) \geq p_{i,i_j}(md)p_{i_j,i_{j+1}} \cdots p_{i_{k-1},i_k}p_{i_k,i}(nd) > 0,
\]

which contradicts that the period of \( i \) is \( d \), since \( md + nd \) is divisible by \( d \) but \( k - j \) is not. Therefore, no such state \( i \) exists and we conclude that the classes \( C_0, \ldots, C_{d-1} \) are mutually disjoint.

(c) (3 marks) Clearly, state \( i_0 \) belongs to at least one of \( C_0, \ldots, C_{d-1} \) (it belongs to \( C_0 \)). Let \( i \) be any given state not equal to \( i_0 \). Since states \( i \) and \( i_0 \) communicate there is some \( m \) such that \( p_{i,i_0}(m) > 0 \). Then for some \( k \in \{0, 1, \ldots, d - 1\} \), we have that \( m + k \) is divisible by \( d \), and so there is a path segment from state \( i \) to state \( i_k \) whose length is divisible by \( d \), since

\[
p_{i,i_k}(m + k) \geq p_{i,i_0}(m)p_{i_0,i_1} \cdots p_{i_{k-1},i_k} > 0.
\]

By part(a), state \( i \) belongs to class \( C_k \).

(d) (3 marks) Suppose that the Markov chain starts in state \( i \in C_j \). Suppose it were possible to revisit class \( C_j \) at a time \( m \) that was not a multiple of \( d \), i.e., for some state \( k \in C_j \) we have \( p_{i,k}(m) > 0 \) with \( m \) not a multiple of \( d \). Since \( k \in C_j \), by part(a) there is some \( n_0 \) such that \( p_{k,i_j}(n_0d) > 0 \), and since \( i \in C_j \), by definition there is some \( n \) such that \( p_{i,j}(nd) > 0 \). So then

\[
p_{i}(m + n_0d + nd) \geq p_{i,k}(m)p_{k,i_j}(n_0d)p_{i,j}(nd) > 0.
\]

But this contradicts the period of state \( i \) being \( d \) since \( m + n_0d + nd \) is not divisible by \( d \). Therefore, starting in class \( C_j \), it is not possible to revisit class \( C_j \) except at times that are a multiple of \( d \). This then implies that the process must cycle through the classes, in some fixed order. To see this start the process in some initial class. In the next \( d - 1 \) transitions, the process must visit a previously unvisited class, so clearly the process must visit the other \( d - 1 \) classes, in some order. On the \( d \)th transition there are no unvisited classes left to visit and so the only class the process could return to is the initial class. On the transition after this the process has no choice but to visit the class it visited on its first transition, and so on. Finally, the order in which the classes are visited must be the order given in the problem (i.e., \( C_0 \rightarrow C_1 \rightarrow \ldots \rightarrow C_{d-1} \rightarrow C_0 \ldots \)). For suppose that the process was able to cycle through the classes in some other order.
Regardless of what state the process starts in, by irreducibility there is some finite sequence of transitions with positive probability that takes the process to state \( i_0 \). But from state \( i_0 \) we know we can cycle through the classes in the order \( C_0 \rightarrow C_1 \rightarrow \ldots \rightarrow C_{d-1} \). Thus, it becomes possible to start cycling through the classes in some order and then to change the order. This contradicts our previous discussion.

\*6. From Sheet. (6 marks)

(a) (3 marks) If \( X = \{X_n : n \geq 0\} \) and \( Y = \{Y_n : n \geq 0\} \) are Markov chains then the sequence \( Z = \{Z_n = X_n + Y_n : n \geq 0\} \) is not necessarily a Markov chain. For example, let \( X \) and \( Y \) be two independent Markov chains with state spaces \( S_X = \{0, 1\} \) and \( S_Y = \{1, 2\} \) respectively, and let the transition probability matrices be as follows:

\[
P_X = \begin{bmatrix} 1 & 0 \\ 3/4 & 1/4 \end{bmatrix} \quad \text{and} \quad P_Y = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.
\]

Note that state 0 is an absorbing state for the \( X \) process, and the state space of the \( Z \) process is \( \{1, 2, 3\} \). We can easily confirm that

\[
P(Z_n = 3 \mid Z_{n-1} = 2) = \sum_{k=1}^{3} P(Z_n = 3 \mid Z_{n-1} = 2, Z_{n-2} = k)P(Z_{n-2} = k) > 0,
\]

since one can readily verify that both \( P(Z_n = 3 \mid Z_{n-1} = 2, Z_{n-2} = 3) \) and \( P(Z_{n-2} = 3) \) are both positive. On the other hand,

\[
P(Z_n = 3 \mid Z_{n-1} = 2, Z_{n-2} = 1) = 0,
\]

since knowing \( Z_{n-2} = 1 \) implies that the \( X \) process is in state 0 so that the \( Z \) process could never go into state 3.

As another example, suppose \( \{X_n : n \geq 0\} \) is a simple random walk with \( p_{i,i+1} = p \) and \( p_{i,i-1} = 1 - p \), for every integer \( i \), and let \( Y_n = \max\{X_k : 0 \leq k \leq n\} - X_n \). One can verify that \( \{Y_n : n \geq 0\} \) is a Markov chain with transition probabilities \( p_{i,i+1} = 1 - p \) and \( p_{i,i-1} = p \) if \( i \geq 1 \), and \( p_{0,0} = p = 1 - p_{0,1} \). One can also readily verify that the sum \( Z_n = \max\{X_k : 0 \leq k \leq n\} \) is not a Markov chain.

(b) (3 marks) For \( n \geq 1 \), consider the sequence

\[
a_i(n) = \begin{cases} 1/n & \text{for } i = 0, \ldots, n-1 \\ 1/2^{i-n+1} & \text{for } i \geq n \end{cases}
\]
Then 
\[ \sum_{i=0}^{\infty} a_i(n) = \sum_{i=0}^{n-1} \frac{1}{n} + \sum_{i=n}^{\infty} \frac{1}{2^{i-n+1}} = 1 + \sum_{i=1}^{\infty} \frac{1}{2^i} = 1 + 1 = 2. \]

Therefore, 
\[ \lim_{n \to \infty} \sum_{i=0}^{\infty} a_i(n) = \lim_{n \to \infty} 2 = 2. \]

However, for a given \( i \), 
\[ a_i(n) = \begin{cases} 
\frac{1}{2^{i-n+1}} & \text{for } n \leq i \\
\frac{1}{n} & \text{for } n > i,
\end{cases} \]

and so 
\[ \sum_{i=0}^{\infty} \lim_{n \to \infty} a_i(n) = \sum_{i=0}^{\infty} \lim_{n \to \infty} \frac{1}{n} = \sum_{i=0}^{\infty} (0) = 0. \]