1. Consider a Bernoulli process \( (X_n : n \in \mathbb{N}) \). Let \( N \) be the first instance where there are two consecutive successes, i.e.,
\[
N = \inf\{ n : X_{n-1} = X_n = 1 \}.
\]
What is the probability \( P(X_{N+1} = X_{N+2} = 0) \) that the next two trials result in failures?

**Solution:** Observe that \( N \) is a stopping time and \( E[N] < \infty \), which implies that \( P(N < \infty) = 1 \). Therefore,
\[
P(X_{N+1} = X_{N+2} = 0) = \sum_{n \in \mathbb{N}} P(X_{N+1} = 0, X_{N+2} = 0, N = n)
\]
\[
= \sum_{n \in \mathbb{N}} P(N = n) \cdot P(X_{n+1} = 0 | N = n) \cdot P(X_{n+2} = 0 | X_{n+1} = 0, N = n).
\]
We now observe that the event \( \{ N = n \} \) is a function only of \( X_1, \ldots, X_n \) as a virtue of the fact that \( N \) is a stopping time. Hence, the second term of the product on the right hand side of the above equation is equal to the unconditional probability of \( \{X_{n+1} = 0\} \) and the third term of the product on the right hand side of the above equation is equal to the unconditional probability of \( \{X_{n+2} = 0\} \), since \( X_{n+1} \) is independent of \( X_1, \ldots, X_n \) and \( X_{n+2} \) is independent of \( X_1 \ldots, X_{n+1} \) respectively. Hence, we get
\[
P(X_{N+1} = X_{N+2} = 0) = \sum_{n \in \mathbb{N}} P(N = n) \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.
\]

2. Let \( (X_n)_{n \geq 0} \) be a discrete-time Markov chain on a countable state space \( S \). Prove the following statements:

(a) For any subsets \( A_0, \ldots, A_{n-1} \subset S \), and for any \( i, j \in S \), we have
\[
P(X_{n+1} = j | X_0 = A_0, \ldots, X_{n-1} = A_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i).
\]

(b) For \( n < n_1 < n_2 \), and any \( i_0, \ldots, i_{n-1}, i, j_1, j_2 \in S \), we have
\[
P(X_{n_1} = j_1, X_{n_2} = j_2 | X_0 = i_0, \ldots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n_1} = j_1, X_{n_2} = j_2 | X_n = i).
\]

**Solution:**

(a) We have
\[
P(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}, X_n = i, X_{n+1} = j)
\]
\[
= \sum_{l_0, \ldots, l_{n-1} \in S} P(X_0 = l_0, \ldots, X_{n-1} = l_{n-1}, X_n = i, X_{n+1} = j)
\]
\[
= \sum_{l_0, \ldots, l_{n-1} \in S} P(X_0 = l_0, \ldots, X_{n-1} = l_{n-1}) \cdot P(X_{n+1} = j | X_0 = l_0, \ldots, X_{n-1} = l_{n-1})
\]
\[
= \sum_{l_0, \ldots, l_{n-1} \in S} P(X_0 = l_0, \ldots, X_{n-1} = l_{n-1}) \cdot P(X_{n+1} = j | X_n = i)
\]
\[
= P(X_{n+1} = j | X_n = i) \cdot P(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}).
\]
Dividing both sides of the above equation by \( P(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}) \) yields the desired result.

(b) This part may be proved along the lines of the steps shown in part (a) above by introducing all the intermediate states the DTMC visits between times \( n \) and \( n_1 \), and between times \( n_1 \) and \( n_2 \).
3. Let $X = (X_n : n \in \mathbb{N}_0)$ be a random process taking values in $S$, a discrete set. Show that $X$ is a time homogeneous Markov chain if and only if there exists a function $f(i, j)$ such that
\[
P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) = f(i, j).
\]
holds for all $i_0, i_1, \ldots, i_{n-1}, i, j \in S$.

**Solution:** The “only if” part is shown as follows. If $(X_n)_{n \geq 0}$ is a time homogeneous DTMC, then,
\[
P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_n = i) = P(X_{n+1} = j | X_n = i)
\]
which is a function only of $i$ and $j$. Now, suppose there exists a function $f(i, j)$ such that
\[
P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) = f(i, j).
\]
holds for all $i_0, i_1, \ldots, i_{n-1}, i, j \in S$. Then, we first demonstrate that $(X_n)_{n \geq 0}$ forms a DTMC. Towards this end, we have
\[
P(X_{n+1} = j | X_n = i) = \sum_{i_0, \ldots, i_{n-1} \in S} P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i) \cdot P(X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1})
\]
\[
= f(i, j) \cdot \sum_{i_0, \ldots, i_{n-1} \in S} P(X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1})
\]
\[
= f(i, j) = P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}, X_n = i).
\]

Next, we show that $(X_n)_{n \geq 0}$ is time homogeneous. We notice that the above set of equations demonstrate that $P(X_{n+1} = j | X_n = i) = f(i, j)$ for every time $n \geq 0$. In particular, $f(i, j) = P(X_1 = j | X_0 = i)$. Thus, the DTMC is time homogeneous.

4. There are $N$ empty boxes and an infinite collection of balls. At each step, a box is chosen at random and a ball is placed in it. Let $X_n$ be the number of empty boxes after the $n^{th}$ ball has been placed.

1. Show that $(X_n : n \in \mathbb{N})$ is a Markov chain.
2. Display its state space and transition probabilities.
3. Classify the states as transient, null or positive.

**Solution:**
(a) Observe that
\[
P[X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0] =
\begin{cases}
\text{probability of picking an empty box,} & \text{if } i = j + 1 \\
\text{probability of picking a box which has one/more balls,} & \text{if } i = j + 1 \\
0, & \text{otherwise}
\end{cases}
\]
Thus, $P[X_{n+1} = j | X_n = i, X_{n-1} = x_{n-1}, \ldots, X_0 = x_0]$ is a function of only $i$ and $j$ thereby letting us conclude that $(X_n)$ is a time homogeneous DTMC.

(b) The transition matrix $P$ is defined as:
\[
P(i, j) = \begin{cases}
\frac{N-i}{N}, & \text{if } i = j \\
1 - \frac{N-i}{N}, & \text{if } i = j + 1 \\
0, & \text{otherwise}
\end{cases}
\]
for $N \geq i \geq 0$. 
5. Let \((X_n)_{n \geq 0}\) be a time homogeneous discrete-time Markov chain on a countable state space \(S\), with transition probabilities \(p_{ij}\). Assume that \(p_{ii} < 1\) for all \(i \in S\). Define

\[ T_1 := \inf\{n \geq 1 : X_n \neq X_0\}, \]

and for \(m \geq 1\), define

\[ T_{m+1} := \inf\{n > T_m : X_n \neq X_{T_m}\}. \]

Thus, the \(T_m\)'s are random times at which the DTMC changes its state.

(a) Verify that for each \(m \geq 1\), \(T_m\) is a stopping time, with \(P(T_m < \infty) = 1\).

(b) Define \(Z_0 = X_0\) and \(Z_m = X_{T_m}\) for \(m \geq 1\). Thus, the process \((Z_m)_{m \geq 0}\) is formed by observing the DTMC \((X_n)_{n \geq 1}\) only at the times when it changes state. Prove that \((Z_m)_{m \geq 0}\) is a time homogeneous DTMC with transition probabilities \(\tilde{p}_{ij}\) given by

\[ \tilde{p}_{ij} = P(Z_{m+1} = j | Z_m = i) = \begin{cases} 0, & i = j \\ \frac{p_{ij}}{1 - p_{ii}}, & i \neq j. \end{cases} \]

(c) What is the relation between the invariant pmf of the \((X_n)_{n \geq 0}\) process and the invariant pmf of the \((Z_m)_{m \geq 0}\) process?

**Solution:**

(a) We prove by induction that \(T_m\) is a stopping time for each \(m \geq 1\). For \(m = 1\), we have

\[ \{T_1 = n\} = \{X_1 = X_0, \ldots, X_{n-1} = X_0, X_n \neq X_0\}, \]

which is a function only of \(\{X_0, \ldots, X_n\}\) for each \(n \geq 0\), and hence \(T_1\) is a stopping time. We now assume that \(\{T_k, 0 \leq k \leq m\}\) is a collection of stopping times, where \(T_0 = 0\). Then,

\[ \{T_{m+1} = n\} = \bigcap_{k=1}^{m} \{X_{T_k+1} = X_{T_k}, \ldots, X_{T_{k-1} + 1} = X_{T_{k-1}}, X_{T_{k-1} + 1} \neq X_{T_k}\}, \]

and since each of \(T_0, \ldots, T_m\) should be less than \(n\) when \(T_{m+1} = n\), it follows that the event on the right hand side of the above equation is a function only of \(X_0, \ldots, X_n\), and thus \(T_{m+1}\) is a stopping time. Further, since for each state \(i \in S\) we have \(p_{ii} < 1\), it follows that \(f_{ii} \leq p_{ii} < 1\), and therefore every state is transient. This implies that the DTMC leaves every state \(i\) in finite time with probability 1.

(b) To show that \((Z_m)_{m \geq 0}\) is a DTMC, by the Strong Markov property, we have

\begin{align*}
P(Z_{m+1} = j | Z_m = i, Z_{m-1} = i_{m-1}, \ldots, Z_0 = i_0) &= P(X_{T_{m+1}} = j | X_{T_m} = i, X_{T_{m-1}} = i_{m-1}, \ldots, X_0 = i_0) \\
&= P(X_{T_{m+1}} = j | X_{T_m} = i) \\
&= P(Z_{m+1} = j | Z_m = i).
\end{align*}
In order to show that \((Z_m)_{m \geq 0}\) is time homogeneous, we claim that \(P(X_{T_{m+1}} = j | X_m = i) = P(X_{T_1} = j | X_0 = i)\) when \(j \neq i\) and 0 otherwise. Indeed, by using the fact that \(P(T_m < \infty) = 1\) and \(P(T_{m+1} - T_m < \infty) = 1\), we have

\[
P(X_{T_{m+1}} = j, X_{T_m} = i) = P(X_{T_{m+1}} = j, X_{T_m} = i, T_m < \infty, T_{m+1} - T_m < \infty)
\]

\[
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} P(X_{T_{m+1}} = j, X_{T_m} = i, T_m = k, T_{m+1} - T_m = l)
\]

\[
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} P(X_k = i, T_m = k, X_{k+1} = i, \ldots, X_{k+l-1} = i, X_{k+l} = j)
\]

\[
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} P(T_m = k, X_k = i) \cdot P(X_{k+1} = i, \ldots, X_{k+l-1} = i, X_{k+l} = j | X_k = i, T_m = k) \cdot 1_{i \neq j}
\]

\[
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} P(T_m = k, X_k = i) \cdot P(T_1 = l, X_l = j | X_0 = i) \cdot 1_{i \neq j}
\]

\[
= \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} P(T_m = k, X_k = i) \cdot P(T_1 = l, X_1 = j | X_0 = i) \cdot 1_{i \neq j}
\]

\[
= 1_{j \neq i} P(X_{T_m} = i) \cdot P(X_{T_1} = j | X_0 = i),
\]

where (a) above follows from the Markov property of \((X_n)_{n \geq 0}\) and the fact that \(T_m\) is a stopping time, and (b) follows from time homogeneity of \((X_n)_{n \geq 0}\). Dividing both sides of the last line of the above sequence of equations by \(P(X_{T_m} = i)\) proves our claim. We have thus established that

\[
P(Z_{m+1} = j | Z_m = i) = P(X_{T_1} = j | X_0 = i) \cdot 1_{j \neq i} = P(Z_1 = j | Z_0 = i) \cdot 1_{j \neq i},
\]

thereby demonstrating that \((Z_m)_{m \geq 0}\) is time homogeneous.

(c) Since \((Z_m)_{m \geq 0}\) is time homogeneous, we have

\[
P(Z_{m+1} = j | Z_m = i) = 1_{j \neq i} \cdot P(Z_1 = j | Z_0 = i)
\]

\[
= 1_{i \neq j} \cdot P(X_{T_1} = j | X_0 = i)
\]

\[
= 1_{i \neq j} \cdot P(X_{T_1} = j, T_1 < \infty | X_0 = i)
\]

\[
= 1_{i \neq j} \cdot \sum_{k=1}^{m} P(X_{T_1} = j, T_1 = k | X_0 = i)
\]

\[
= 1_{i \neq j} \cdot \sum_{k=1}^{m} P(X_k = j, X_{k-1} = i, \ldots, X_1 = i | X_0 = i)
\]

\[
= 1_{i \neq j} \cdot \sum_{k=1}^{m} (p_{ii})^{k-1} \cdot p_{ij}
\]

\[
= 1_{i \neq j} \cdot \frac{p_{ij}}{1 - p_{ii}}.
\]

6. Let \((X_n)_{n \geq 0}\) is a stochastic process taking values in a discrete state space \(\mathcal{S}\).

(a) Show that if \((X_n)_{n \geq 0}\) is (strictly) stationary, then for all \(m, n \geq 0\) and for all \(i \in \mathcal{S}\),

\[
P(X_n = i) = P(X_m = i),
\]

i.e., the distribution of \(X_n\) is invariant with \(n\).
7. \( \{X_n, n \geq 0\} \) is an irreducible, positive recurrent DTMC on \( S = \{0, 1, 2, 3, \ldots\} \). Let \( X_0 = j \). Fix an \( i \in S \), \( i \neq j \), and for \( k \geq 1 \), define the random variable \( V_k^{(i)} \) as the number of visits to \( i \) between the \((k-1)^{th}\) and \( k^{th}\) return to \( j \). Show that \( (V_k^{(i)})_{k \geq 1} \) is an iid sequence.

**Solution:**

(a) Since the DTMC is positive recurrent, it will return back to state \( j \) in finite time with probability 1, i.e., \( P(T_k < \infty) = 1 \) for all \( k \geq 1 \). This in turn implies that \( P(T_k - T_{k-1} < \infty) = 1 \) for each \( k \geq 1 \), where \( T_0 = 0 \). In what follows, we first demonstrate that \( V_1^{(i)} \) and \( V_k^{(i)} \) are identically distributed for each \( k \geq 1 \). For any \( n \in \{0, 1, 2, \ldots\} \), we have

\[
P(V_k^{(i)} = n) = P(V_k^{(i)} = n, T_{k-1} < \infty, T_k - T_{k-1} < \infty)
\]

\[
= \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} P(V_k^{(i)} = n, T_{k-1} = m, T_k - T_{k-1} = l)
\]

\[
= \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} P(V_k^{(i)} = n, T_{k-1} = m, X_m = j, T_k - T_{k-1} = l)
\]

\[
= \sum_{m,j=1}^{\infty} P(V_k^{(i)} = n, T_{k-1} = m, X_m = j, x_{m+1} \neq j, \ldots, x_{m+l-1} \neq j, x_{m+l} = j) \quad \text{(says that } T_k - T_{k-1} = l \text{)}
\]

\[
= \sum_{m,l=1}^{\infty} P(T_{k-1} = m, X_m = j) \cdot P(V_k^{(i)} = n, x_{m+1} \neq j, \ldots, x_{m+l-1} \neq j, x_{m+l} = j | X_m = j, T_{k-1} = m)
\]

\[
\overset{(a)}{= \sum_{m,l=1}^{\infty}} P(T_{k-1} = m) \cdot P(V_k^{(i)} = n) \overset{X_1 \neq j, \ldots, X_{l-1} \neq j, X_l = j | X_0 = j}{=} \overset{T_1 = l}{P(V_1^{(i)} = n)}
\]

\[
= \sum_{m,l=1}^{\infty} P(T_{k-1} = m) \cdot P(V_k^{(i)} = n, T_1 = l | X_0 = j)
\]

\[
= P(V_1^{(i)} = n | X_0 = j)
\]

\[
= P(V_1^{(i)} = n),
\]
where the last line above follows since $X_0 = j$ and (a) follows from the fact that the event of hitting state $i$ $n$ times and not visiting state $j$ for $l$ consecutive instants of time starting from time $m$ is identical to the event of hitting state $i$ $n$ times and not visiting state $j$ for $l$ consecutive instants of time starting from time 0 by Markov property.

Next, we use this identical distribution property to show that for each $k \geq 1$, the random variables $V_{k+1}^{(i)}$ and $V_k^{(i)}$ are independent. Similar logic may be applied to show independence of any finite collection. We have

$$P(V_{k+1}^{(i)} = m, V_k^{(i)} = n) = P(V_{k+1}^{(i)} = m, V_k^{(i)} = n, T_{k-1} < \infty, T_k - T_{k-1} < \infty, T_{k+1} - T_k < \infty)$$

$$= \sum_{r,s,t=1}^{\infty} P(V_{k+1}^{(i)} = m, V_k^{(i)} = n, T_{k-1} = r, X_r = j, T_k - T_{k-1} = s, T_{k+1} - T_k = t)$$

$$= \sum_{r=1}^{\infty} P(T_{k-1} = r, X_r = j) \cdot P(V_k^{(i)} = n, T_1 = s | X_0 = j) \cdot P(V_{k+1}^{(i)} = m, T_1 = t)$$

$$= P(V_k^{(i)} = m) \cdot P(V_{k+1}^{(i)} = n)$$

where (a) above follows from the same steps as demonstrated for identical distribution before.

8. $(X_n)_{n \geq 0}$ is a DTMC on $\{0, 1, 2, \ldots\}$, with $p_{0i} = \left(\frac{1}{2}\right)^i$ for $i \in \{1, 2, \ldots\}$, and for $i \geq 1$, $p_{i0} = \frac{1}{2}$ and $p_{i,i+1} = \frac{1}{2}$.

(a) Show the state transition diagram for this DTMC.

(b) Is this DTMC irreducible?

(c) Find $f_{0i}^{(n)}$ for $n \geq 1$.

(d) Hence conclude that the DTMC is positive recurrent.

**Solution:**

(a) The figure below depicts the state transition diagram.

![State Transition Diagram](image)

(b) From the state transition diagram, we observe that between any two states, there exists a path of positive probability. Thus, the DTMC is irreducible.
(c) For \( n \geq 2 \), we have

\[
 f^{(n)}_{00} = \sum_{i=1}^{\infty} \left( \frac{1}{2} \right)^i \cdot \left( \frac{1}{2} \right)^{n-2} \cdot \frac{1}{2},
\]

where the first and last terms in the product on the right hand side of the above equation correspond to the forward step from 0 and the backward return step to 0 respectively, and the intermediate term corresponds to the journey away from 0. We thus have \( f^{(n)}_{00} = \left( \frac{1}{2} \right)^{n-1} \) for all \( n \geq 2 \). Further,

\[
 f_{00} = \sum_{n=2}^{\infty} \left( \frac{1}{2} \right)^{n-1} = 1,
\]

and

\[
 \mu_{00} = \sum_{n=2}^{\infty} n \cdot \left( \frac{1}{2} \right)^{n-1} = 3.
\]

Thus, state 0 is positive recurrent.

(d) Since the DTMC is irreducible, and state 0 is positive recurrent, the DTMC is positive recurrent.