1. Let $X_0, Y_1, Y_2, \ldots$ be independent random variables, $X_0$ with values in $\{0, 1, 2, \ldots \}$ and each $Y_i$ an indicator random variable with $\mathbb{P}(Y_i = 1) = \frac{1}{i}$ and $\mathbb{P}(Y_i = 0) = 1 - \frac{1}{i} = \frac{i-1}{i}$ for each $i = 1, 2, \ldots$ For $n = 1, 2, \ldots$ let

$$X_{n+1} := \begin{cases} \max \{ k : 1 \leq k < X_n \text{ and } Y_k = 1 \} & \text{if } X_n > 1, \\ 0 & \text{if } X_n \leq 1. \end{cases}$$

Explain why $(X_n)$ is a Markov chain, and describe its state space and transition probabilities.

**Solution:** The state space is clearly $\{0, 1, 2, \ldots \}$ and, moreover, $X_{n+1} < X_n$ when $X_n > 1$. Suppose $X_i > 1$ and $0 < X_{i+1} < X_i$ for $i \in \{0, 1, 2, \ldots, n\}$. Then

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_i = x_i \text{ for } i = 0, 1, \ldots, n) = \frac{\mathbb{P}(X_{i+1} = x_{i+1} \text{ for } i = 0, 1, 2, \ldots, n+1)}{\mathbb{P}(X_{i+1} = x_{i+1} \text{ for } i = 0, 1, 2, \ldots, n)} = \frac{\mathbb{P}(Y_{x_i+1} = 1, Y_{x_i+1+1} = \cdot \cdot \cdot = Y_{x_i+1+t} = 0, Y_{x_i+1+t+1} = 1)}{\mathbb{P}(Y_{x_i} = 1) \prod_{i=1}^{x_i+1} \{ \mathbb{P}(Y_j = 0) \mathbb{P}(Y_{x_i} = 1) \}}.$$

Many numerator/denominator cancellations occur and all that remains after cancellations is one term of the numerator’s outer big-product:

$$\left( \prod_{j=x_{n+1}+1}^{x_{n+1}+1} \mathbb{P}(Y_j = 0) \right) \frac{\mathbb{P}(Y_{x_{n+1}+1} = 1)}{\mathbb{P}(Y_{x_{n+1}} = 1)} = \frac{1}{x_{n+1}} \prod_{j=x_{n+1}+1}^{x_{n+1}+1} \frac{j-1}{j} = \frac{1}{x_{n+1}} \left( \frac{x_{n+1}+1}{x_{n+1}+1+2 \cdot x_{n+1}+1} \right) = \frac{1}{x_{n+1}+1}.$$

Conclude that

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_i = x_i \text{ for } i = 0, 1, \ldots, n) = \frac{1}{x_{n+1}+1}. \quad (1)$$

The above result holds for all $n$ such that $X_i > 1$ and $0 < X_{i+1} < X_i$ for all $0 \leq i \leq n$. The only other case is if there is an $m$ such that $X_m \leq 1$. Note that by how $(X_n)$ is defined, we must have $X_{m+1} = 0$ and trivially we have, for all $n$,

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = 0) = \mathbb{1}(x_{n+1} = 0). \quad (2)$$

Therefore, combining both cases (Equations (1) and (2)), we have

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_i = x_i \text{ for } i = 0, 1, \ldots, n) = \begin{cases} \frac{1}{x_{n+1}} & \text{if } x_n > 1 \text{ and } 0 < x_{n+1} < x_n, \\ \mathbb{1}(x_{n+1} = 0) & \text{if } x_n \leq 1, \\ 0 & \text{otherwise}. \end{cases} \quad (3)$$

In particular, $\mathbb{P}(X_{n+1} = x_{n+1} | X_i = x_i \text{ for } i = 0, 1, \ldots, n)$ does not depend on $x_0, x_1, \ldots, x_{n-1}$, so $\mathbb{P}(X_{n+1} | X_0, \ldots, X_n) = \mathbb{P}(X_{n+1} \mid X_n)$, i.e. $(X_n)$ is a Markov chain with transition probabilities given by Equation (3).

2. For $Y_1, Y_2, \ldots$ as in the previous question, let $T_0 := 0$ and for $n = 1, 2, \ldots$ let

$$T_n := \min \{ k : k > T_{n-1} \text{ and } Y_k = 1 \}.$$

Explain why $(T_n)$ is a Markov chain, and describe its state space and transition probabilities.
Solution: The state space is clearly \( \{0, 1, 2, \ldots \} \) and, moreover, \( T_{n+1} > T_n \) for all \( n \). Note that \( \mathbb{P}(T_1 = 1 \mid T_0 = 0) = 1 \) since \( Y_1 = 1 \) with probability 1. Consider \( n \geq 2 \). We have for \( t_{n+1} > t_n > t_{n-1} > \cdots > t_2 > 1 \):

\[
\mathbb{P}(T_{n+1} = t_{n+1} \mid T_0 = 0, T_1 = 1, T_2 = t_2, \ldots, T_n = t_n) = \frac{\mathbb{P}(T_0 = 0, T_1 = 1, T_2 = t_2, \ldots, T_n = t_n, T_{n+1} = t_{n+1})}{\mathbb{P}(T_0 = 0, T_1 = 1, T_2 = t_2, \ldots, T_n = t_n)} = \frac{\mathbb{P}(T_0 = 0) \mathbb{P}(T_1 = 1 \mid T_0 = 0) \prod_{i=2}^{n+1} \left\{ \left( \prod_{j=t_i-1}^{t_{i+1}-1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_i = 1) \right\}}{\mathbb{P}(T_0 = 0) \mathbb{P}(T_1 = 1 \mid T_0 = 0) \prod_{i=2}^{n+1} \left\{ \left( \prod_{j=t_i-1}^{t_{i+1}-1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_i = 1) \right\}}.
\]

Many numerator/denominator cancellations occur and all that remains after cancellations is one term of the numerator’s outer big-product:

\[
\left( \prod_{j=t_{n+1}-1}^{t_{n+1}} \mathbb{P}(Y_j = 0) \right) \frac{\mathbb{P}(Y_{t_{n+1}} = 1)}{\mathbb{P}(Y_{t_{n+1}} = 1)} = \frac{1}{t_{n+1}} \left( \prod_{j=t_{n+1}}^{t_{n+1}-1} \left( j - 1 \right) \right) = \frac{1}{t_{n+1}} \left( \frac{t_n}{t_n + 1} \frac{t_{n+1} + 1}{t_{n+1} + 2} \cdots \frac{t_{n+1} - 2}{t_{n+1} - 1} \right) = \frac{t_n}{t_{n+1}(t_{n+1} - 1)}.
\]

Conclude that for \( n \geq 2 \),

\[
\mathbb{P}(T_{n+1} = t_{n+1} \mid T_0 = 0, T_1 = 1, T_2 = t_2, \ldots, T_n = t_n) = \begin{cases} \frac{t_n}{t_{n+1}(t_{n+1} - 1)} & \text{if } t_{n+1} > t_n, \\ 0 & \text{otherwise}. \end{cases} \tag{4}
\]

In particular, \( \mathbb{P}(T_{n+1} = t_{n+1} \mid T_i = t_i \text{ for } i = 0, 1, \ldots, n) \) does not depend on \( t_0, t_1, \ldots, t_{n-1} \), so \( \mathbb{P}(T_{n+1} \mid T_0, \ldots, T_n) = \mathbb{P}(T_{n+1} \mid T_n) \), i.e. \( T_n \) is a Markov chain with transition probabilities given by Equation (4).

3. Let \( X, Y, Z \) be random variables defined on a common probability space, each with a discrete distribution. Explain why the function \( \phi(x) := \mathbb{E}(Y \mid X = x) \) is characterized by the property

\[
\mathbb{E}(Y \phi(X)) = \mathbb{E}[\phi(X) g(X)] \tag{5}
\]

for every bounded function \( g \) whose domain is the range of \( X \). Use this characterization of \( \mathbb{E}(Y \mid X) \) to verify the formula

\[
\mathbb{E}(\mathbb{E}(Y \mid X) \mid f(X)) = \mathbb{E}[Y \mid f(X)] \tag{6}
\]

for every function \( f \) whose domain is the range of \( X \), and the formula

\[
\mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid X) = \mathbb{E}[Y \mid X]. \tag{7}
\]

Solution: We first show that \( \phi(x) = \mathbb{E}(Y \mid X = x) \) satisfies Equation (5):

\[
\mathbb{E}(Y \phi(X)) = \sum_x \mathbb{P}(X = x) \mathbb{E}(Y \phi(X) \mid X = x) = \sum_x \mathbb{P}(X = x) g(x) \mathbb{E}(Y \mid X = x) \mathbb{E}(\phi(X) \mid X = x) = \mathbb{E}(g(X) \phi(X)).
\]

Next we show that \( \phi \) is unique, i.e. if a function \( \phi \) satisfies Equation (5), then we must have \( \phi(x) = \mathbb{E}(Y \mid X = x) \). Note that the domain of \( \phi \) is \( \{x : \mathbb{P}(X = x) > 0\} \). Let \( x \in \{x : \mathbb{P}(X = x) > 0\} \). To see that \( \phi(x) \) must be equal to \( \mathbb{E}(Y \mid X = x) \), by Equation (5), we have

\[
\mathbb{E}(Y \mathbf{1}(X = x)) = \mathbb{E}(\phi(X) \mathbf{1}(X = x)) = \phi(x) \mathbb{P}(X = x).
\]

This implies that

\[
\phi(x) = \frac{\mathbb{E}(Y \mathbf{1}(X = x))}{\mathbb{P}(X = x)} = \mathbb{E}(Y \mid X),
\]

using the identity that \( \mathbb{E}(A \mid B) = \mathbb{E}(A\mathbf{1}_B) / \mathbb{P}(B) \). To verify Equation (6), observe that

\[
\mathbb{E}(\mathbb{E}(Y \mid X) \mid f(X)) = \mathbb{E}(\phi(X) \mid f(X)) = \mathbb{E}(\phi(X) \mid f(X)) \mathbb{P}(f(X) = x) = \mathbb{E}(Y \mathbf{1}(f(X) = f(x))) / \mathbb{P}(f(X) = x) = \mathbb{E}(Y \mid f(X) = f(x)).
\]
4. Suppose that a sequence of random variables \( X_0, X_1, \ldots \) and a function \( f \) are such that

\[
\mathbb{E}(f(X_{n+1}) \mid X_0, \ldots, X_n) = f(X_n) \tag{8}
\]

for every \( n = 0, 1, 2, \ldots \). Explain why this implies

\[
\mathbb{E}(f(X_{n+1}) \mid f(X_0), \ldots, f(X_n)) = f(X_n) \tag{9}
\]

Give an example of such an \( f \) which is not constant for \( (X_n) \) a \( p \uparrow, 1 - p \downarrow \) random walk on the integers.

**Solution:** Define random vectors \( \mathbf{X}^{(n)} = (X_0 \ X_1 \ \cdots \ X_{n-1})^\top \) and \( \mathbf{Y}^{(n)} = (f(X_n) \ 0 \ \cdots \ 0)^\top \) taking on values in \( \mathbb{R}^n \). Define function \( g \) by \( g(\mathbf{X}^{(n)}) = (f(X_0) \ f(X_1) \ \cdots \ f(X_{n-1}))^\top \). Then

\[
\mathbb{E}(f(X_n) \mid f(X_0), \ldots, f(X_{n-1})) = \mathbb{E}\left( \begin{pmatrix} 1 & 0 & \cdots & 0 \\ f(X_0) & f(X_1) & \cdots & f(X_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \mid \begin{pmatrix} f(X_0) \\ f(X_1) \\ \vdots \\ f(X_{n-1}) \end{pmatrix} \right)
\]

\[
= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ f(X_0) & f(X_1) & \cdots & f(X_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\mathbb{E}(\mathbf{Y}^{(n)} \mid \mathbf{X}^{(n)})
\]

\[
= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ f(X_0) & f(X_1) & \cdots & f(X_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\mathbb{E}(\mathbf{Y}^{(n)} \mid \mathbf{X}^{(n)}) \quad \text{by Equation (6)}
\]

\[
= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ f(X_0) & f(X_1) & \cdots & f(X_{n-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}\mathbb{E}(\mathbf{Y}^{(n)} \mid \mathbf{X}^{(n)}) \quad \text{by Equation (8)}
\]

\[
= \mathbb{E}(f(X_{n-1}) \mid f(X_0), \ldots, f(X_{n-1})),
\]

We can verify Equation (7) with direct computation:

\[
\mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid X = x) = \sum_z \mathbb{E}(Y \mid X = x, Z = z) \mathbb{P}(Z = z \mid X = x)
\]

\[
= \sum_z \sum_y y \mathbb{P}(Y = y \mid X = x, Z = z) \mathbb{P}(Z = z \mid X = x)
\]

\[
= \sum_z \sum_y \frac{\mathbb{P}(X = x, Y = y, Z = z)}{\mathbb{P}(X = x, Z = z)} \frac{\mathbb{P}(X = x, Z = z)}{\mathbb{P}(X = x)}
\]

\[
= \sum_z \sum_y \frac{\mathbb{P}(X = x, Y = y, Z = z)}{\mathbb{P}(X = x)}
\]

\[
= \sum_z \sum_y y \mathbb{P}(Y = y \mid X = x)
\]

\[
= \mathbb{E}(Y \mid X = x).
\]
which is precisely Equation (9).

As an example, if \( f(x) = \left( \frac{q}{p} \right)^x \), then if \( (X_n) \) is a \( p \uparrow, 1-p \downarrow \) walk on the integers, then \( (f(X_n)) \) is a martingale since

\[
\mathbb{E}(f(X_{n+1}) \mid X_0, \ldots, X_n) = \mathbb{E}\left( \left( \frac{q}{p} \right)^{X_{n+1}} \mid X_0, \ldots, X_n \right)
\]

\[
= p \left( \frac{q}{p} \right)^{X_{n+1}} + q \left( \frac{q}{p} \right)^{X_n-1}
\]

\[
= \frac{q^{X_n+1}}{p^{X_n}} + \frac{q^{X_n}}{p^{X_n-1}}
\]

\[
= \frac{q^{X_n+1}}{p^{X_n}} + \frac{q^{X_n}p}{p^{X_n}}
\]

\[
= \frac{q^{X_n}q + q^{X_n}p}{p^{X_n}}
\]

\[
= \frac{q^{X_n}}{p^{X_n}} (q + p)
\]

\[
= f(X_n),
\]

so by the result above, we have \( \mathbb{E}(f(X_{n+1}) \mid f(X_0), \ldots, f(X_n)) = f(X_n) \).

5. Let \( S := X_1 + \cdots + X_N \) be the number of successes and \( F := N - S \) be the number of failures in a Poisson (\( \mu \)) distributed random number \( N \) of Bernoulli trials, where given \( N = n \) the \( X_1, \ldots, X_n \) are independent with \( \mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = p \) for some \( 0 \leq p \leq 1 \). Derive the joint distribution of \( S \) and \( F \). How can the conclusion be generalized to multinomial trials?

**Solution:** Let \( q = 1 - p \). We have

\[
\mathbb{P}(S = s, F = f) = \sum_{n=0}^{\infty} \mathbb{P}(S = s, F = f \mid N = n) \mathbb{P}(N = n)
\]

\[
= \sum_{n=0}^{\infty} \mathbb{P}(S = s, N - S = f \mid N = n) \mathbb{P}(N = n)
\]

\[
= \sum_{n=0}^{\infty} \mathbb{P}(S = s, S = N - f \mid N = n) \mathbb{P}(N = n)
\]

\[
= \sum_{n=0}^{\infty} \mathbb{P} \left( \sum_{i=1}^{n} X_i = s, \sum_{i=1}^{n} X_i = n - f \right) \mathbb{P}(N = n)
\]

\[
= \sum_{n=0}^{\infty} 1(s = n - f) \mathbb{P} \left( \sum_{i=1}^{n} X_i = s \right) \mathbb{P}(N = n)
\]

\[
= \sum_{n=0}^{\infty} 1(n = s + f) \mathbb{P} \left( \sum_{i=1}^{n} X_i = s \right) \mathbb{P}(N = n)
\]

\[
= \mathbb{P} \left( \sum_{i=1}^{s+f} X_i = s \right) \mathbb{P}(N = s + f)
\]

\[
= \frac{(s+f)!}{s!} p^s q^f \mu^{s+f} e^{-\mu} (s+f)^f
\]

\[
= \frac{(s+f)!}{s!} p^s q^f \mu^{s+f} e^{-\mu} (s+f)^f
\]

\[
= \mathbb{P} \text{ (Poisson } (p\mu) = s) \mathbb{P} \text{ (Poisson } (q\mu) = f). \]

In the multinomial case with \( k \) categories with probabilities \( p_1, p_2, \ldots, p_k \) and \( N \sim \text{ Poisson } (\mu) \) trials, let \( S_1, S_2, \ldots, S_k \)
denote the number of trials falling into categories 1, 2, \ldots, k respectively. Then generalizing the result above, we have
\[
P(S_1 = s_1, S_2 = s_2, \ldots, S_k = s_k) = \prod_{i=1}^{k} P(\text{Poisson}(p_i \mu) = s_i).
\]

6. Let \( \mathbb{P}_{i} \) govern a \( p \uparrow, q = 1 - p \downarrow \) walk \((S_n)\) on the integers started at \( S_0 = i \), with \( p > q \). Let
\[
f_{ij} := \mathbb{P}_{i}(S_n = j \text{ for some } n \geq 1).
\]

Use results derived from lectures and/or the text to present a formula for \( f_{ij} \) in each of the two cases \( i > j \) and \( i < j \).

**Solution:**

Case \( i > j \):

This can be viewed as the gambler’s ruin problem for a biased coin where the bottom “absorbing” state is \( j \) and the top “absorbing” state is \( +\infty \). \( f_{ij} \) is the probability of starting at \( i \) and hitting \( j \) before hitting \( +\infty \). Using a result from lecture, we have
\[
f_{ij} = \mathbb{P}_{i} \text{ (hit } j \text{ before } +\infty) = \lim_{b \to +\infty} \mathbb{P}_{a} \text{ (hit 0 before } b) = \left(\frac{q}{p}\right)^{a - b},
\]
where \( a = i - j \) and \( b \to +\infty \).

Case \( i < j \):

Claim: Since \( p > q \), we are guaranteed to hit \( j \) starting from \( i \), so \( f_{ij} = \mathbb{P}_{i} \text{ (hit } j) = 1 \). To show this, consider the gambler’s ruin problem where we flip the walk upside down, i.e. suppose we start at \(-i\) and want to reach \(-j\) before we reach \( +\infty \) where a step up has probability \( q \) and a step down has probability \( p \), where \( p > q \). Using the result from class, we have
\[
f_{ij} = \lim_{b \to +\infty} \mathbb{P}_{a} \text{ (hit 0 before } b) = \lim_{b \to +\infty} \left(1 - \frac{\left(\frac{q}{p}\right)^{a - 1}}{\left(\frac{q}{p}\right)^{b - 1}}\right) = 1 - \lim_{b \to +\infty} \frac{\left(\frac{q}{p}\right)^{a - 1} - 1}{\left(\frac{q}{p}\right)^{b - 1} - 1}.
\]

Since \( p > q \), the right-most term’s denominator goes to \(+\infty\) whereas the numerator is fixed, so \( \lim_{b \to +\infty} \frac{\left(\frac{q}{p}\right)^{a - 1} - 1}{\left(\frac{q}{p}\right)^{b - 1} - 1} = 0 \).

Thus, we have \( f_{ij} = 1 - \lim_{b \to +\infty} \frac{\left(\frac{q}{p}\right)^{a - 1} - 1}{\left(\frac{q}{p}\right)^{b - 1} - 1} = 1 - 0 = 1 \).

Case \( i = j \):

From first-step analysis, we have
\[
f_{ii} = \mathbb{P}(\text{go } 1 \text{ step up}) \mathbb{P}_{i+1} \text{ (hit } i \text{ before } +\infty) + \mathbb{P}(\text{go } 1 \text{ step down}) \mathbb{P}_{i-1} \text{ (hit } i) = p \left(\frac{q}{p}\right)^{(i+1) - i} + q \cdot 1 = p \left(\frac{q}{p}\right) + q = 2q.
\]

7. Let \( \mathbb{P}_{i} \) govern \((X_n)\) as a Markov chain starting from \( X_0 = i \), with finite state space \( S \) and transition matrix \( P \) which has a set of absorbing states \( B \). Let \( T := \min \{ n \geq 1 : X_n \in B \} \) and assume that \( \mathbb{P}_{i}(T < \infty) = 1 \) for all \( i \). Derive a formula for
\[
\mathbb{P}_{i}(X_{T-1} = j, X_T = k) \text{ for } i, j \in B^c \text{ and } k \in B
\]
in terms of matrices \( W := (I - Q)^{-1} \) and \( R \), where \( Q \) is the restriction of \( P \) to \( B^c \times B^c \) and \( R \) is the restriction of \( P \) to \( B^c \times B \).

**Solution:**

\[
\mathbb{P}_{i}(X_{T-1} = j, X_T = k) = \sum_{n=1}^{\infty} \mathbb{P}_{i}(X_{T-1} = j, X_T = k, T = n) = \sum_{n=1}^{\infty} P^{n-1}(i, j) P(j, k) = \left(\sum_{m=0}^{\infty} P^{m}(i, j)\right) P(j, k)
\]
\[
W_{ij} = \left( \sum_{m=0}^{\infty} Q^m(i,j) \right) R(j,k) = W(i,j) R(j,k).
\]

8. In the same setting, let \( f_{ij} := P_i(X_n = j \text{ for some } n \geq 1) \). For \( i, j \in B^c \), find and explain a formula for \( f_{ij} \) in terms of \( W_{ij} \) and \( W_{jj} \).

**Solution:** Let \( N_j \) be the total number of times we visit state \( j \) before absorption. Recall that \( W_{ij} = E_i(N_j) \) and \( W_{jj} = E_j(N_j) \). Reaching \( X_n = j \) for some \( n \geq 1 \) is equivalent to saying that there exists a first time that we reach \( j \); thus:

\[
f_{ij} = P_i(X_n = j \text{ for some } n \geq 1) = P(\text{we reach } j \text{ for the first time}).
\]

From first-step analysis:

\[
E_i(N_j) = P(\text{we reach state } j \text{ for the first time}) \cdot E_j(N_j) + P(\text{we never reach state } j) \cdot 0.
\]

Hence, we have

\[
f_{ij} = P(\text{we reach state } j \text{ for the first time}) = \frac{E_i(N_j)}{E_j(N_j)} = \frac{W_{ij}}{W_{jj}}.
\]

9. In the same setting, let \( \phi_i(s) \) denote the probability generating function of \( T \) for the Markov chain started at state \( i \). Derive a system of equations which could be used to determine \( \phi_i(s) \) for all \( i \in S \).

**Solution:** Note that for \( i \in B, P_i(T = 0) = 1 \), i.e. \( \phi_i(s) = 1 \) for \( i \in B \). For \( i \notin B \), clearly \( P_i(T = 0) = 0 \) and for \( n \geq 1 \), use first-step analysis to get

\[
P_i(T = n) = \sum_j P(i,j) P_j(T = n - 1)
\]

\[
= \sum_{j \in B^c} Q(i,j) P_j(T = n - 1) + \sum_{k \in B} R(i,k) P_k(T = n - 1)
\]

\[
= \sum_{j \in B^c} Q(i,j) P_j(T = n - 1) + \sum_{k \in B} R(i,k) \mathbf{1}(n - 1 = 0)
\]

\[
= \sum_{j \in B^c} Q(i,j) P_j(T = n - 1) + \mathbf{1}(n = 1) \sum_{k \in B} R(i,k).
\]

So

\[
\phi_i(s) = P_i(T = 0) + \sum_{n=1}^{\infty} P_i(T = n) s^n
\]

\[
= \sum_{n=1}^{\infty} \left( \sum_{j \in B^c} Q(i,j) P_j(T = n - 1) + \mathbf{1}(n = 1) \sum_{k \in B} R(i,k) \right) s^n
\]

\[
= \sum_{n=1}^{\infty} \sum_{j \in B^c} Q(i,j) P_j(T = n - 1) s^n + \sum_{n=1}^{\infty} \mathbf{1}(n = 1) \sum_{k \in B} R(i,k) s^n
\]

\[
= \sum_{j \in B^c} Q(i,j) \sum_{n=1}^{\infty} P_j(T = n - 1) s^n + \sum_{k \in B} R(i,k) s
\]

\[
= \sum_{j \in B^c} Q(i,j) \sum_{m=0}^{\infty} P_j(T = m) s^{m+1} + \sum_{k \in B} R(i,k) s
\]

\[
= \sum_{j \in B^c} Q(i,j) s \sum_{m=0}^{\infty} P_j(T = m) s^m + \sum_{k \in B} R(i,k) s
\]

\[
= s \sum_{j \in B^c} Q(i,j) \sum_{m=0}^{\infty} P_j(T = m) s^m + s \sum_{k \in B} R(i,k).
\]
= s \sum_{j \in B^c} Q(j, j) \phi_j(s) + s \sum_{k \in B} R(i, k)
= s \sum_{j \in B^c \setminus \{i\}} Q(j, j) \phi_j(s) + sQ(i, i) \phi_i(s) + s \sum_{k \in B} R(i, k).

Rearranging terms gives
\[ s \sum_{j \in B^c \setminus \{i\}} Q(j, j) \phi_j(s) + (sQ(i, i) - 1) \phi_i(s) + s \sum_{k \in B} R(i, k) = 0, \quad \text{for } i \in B^c. \]

10. Let \( X \) be a non-negative integer valued random variable with probability generating function \( \phi(s) \) for \( 0 \leq s \leq 1 \). Let \( N \) be independent of \( X \) with the Geometric \((p)\) distribution \( \mathbb{P}(N = n) = (1 - p)^n p \) for \( n = 0, 1, 2, \ldots \) where \( 0 < p < 1 \). Find a formula for \( \mathbb{P}(N < X) \) in terms of \( \phi \) and \( p \).

\textbf{Solution:}
\[
\mathbb{P}(N < X) = \sum_{x=0}^{\infty} \mathbb{P}(N < X \mid X = x) \mathbb{P}(X = x)
= \sum_{x=0}^{\infty} \mathbb{P}(N < x) \mathbb{P}(X = x)
= \sum_{x=0}^{\infty} \mathbb{P}(N \leq x - 1) \mathbb{P}(X = x)
= \sum_{x=0}^{\infty} (1 - (1 - p)^x) \mathbb{P}(X = x)
= \sum_{x=0}^{\infty} \mathbb{P}(X = x) - \sum_{x=0}^{\infty} (1 - p)^x \mathbb{P}(X = x)
= 1 - \phi(1 - p).
\]

11. Let \( X \) be a non-negative random variable with usual probability generating function \( \phi(s) \) for \( 0 \leq s \leq 1 \). Define the tail probability generating function \( \tau(s) \) by
\[
\tau(s) := \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) s^n.
\]

Use the identity
\[
\mathbb{P}(X = n) = \mathbb{P}(X \geq n) - \mathbb{P}(X \geq n + 1)
\]
to derive a formula for \( \tau(s) \) in terms of \( s \) and \( \phi(s) \) for \( 0 \leq s \leq 1 \). Discuss what happens for \( s = 1 \).

\textbf{Solution:} We have
\[
\phi(s) = \sum_{n=0}^{\infty} \mathbb{P}(X = n) s^n
= \sum_{n=0}^{\infty} (\mathbb{P}(X \geq n) - \mathbb{P}(X \geq n + 1)) s^n
= \sum_{n=0}^{\infty} \mathbb{P}(X \geq n) s^n - \sum_{n=0}^{\infty} \mathbb{P}(X \geq n + 1) s^n
= \mathbb{P}(X \geq 0) + \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) s^n - \sum_{m=1}^{\infty} \mathbb{P}(X \geq m) s^{m-1}
= \underbrace{\mathbb{P}(X \geq 0)}_{1} + \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) s^n - s^{-1} \sum_{m=1}^{\infty} \mathbb{P}(X \geq m) s^m
= 1 + \tau(s) - s^{-1} \tau(s)
= 1 + \tau(s) (1 - s^{-1}),
\]
13. A branching process with Poisson (\(\lambda\)) offspring distribution started with one individual has extinction probability \(p \) with \(0 < p < 1\). Find a formula for \(\lambda\) in terms of \(p\).

**Solution:** The offspring distribution has PGF

\[
\phi(s) = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} s^n = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda s)^n}{n!} = e^{-\lambda} e^{\lambda s} = e^{\lambda (s-1)}.
\]

The extinction probability \(p\) satisfies \(p = \phi(p) = e^{\lambda(p-1)}\). Taking the log of both sides gives \(\log p = \lambda (p-1)\), so

\[
\lambda = \frac{\log p}{p-1}.
\]

14. Suppose \((X_n)\) is a Markov chain with state space \(\{0,1,\ldots,b\}\) for some positive integer \(b\), with states 0 and \(b\) absorbing and no other absorbing states. Suppose also that \((X_n)\) is a martingale.

Evaluate

\[
\lim_{n \to \infty} P_a(X_n = b)
\]

and explain your answer carefully.

**Solution:** We start at \(X_0 = a\). Since \((X_n)\) is a martingale, \(\mathbb{E}[X_n] = \mathbb{E}[X_0] = a\) for all \(n\). So

\[
a = \mathbb{E}[X_n] = \sum_{i=0}^{b} i P_a(X_n = i) = \sum_{i=1}^{b-1} i P_a(X_n = i) + b P_a(X_n = b) .
\]

**Claim:** From any state \(i \in \{1,2,\ldots,b-1\}\), we can eventually reach an absorbing state with probability 1. Assuming that this claim is true, then for any state \(i \in \{1,2,\ldots,b-1\}\), \(\lim_{n \to \infty} P_a(X_n = i) = 0\). Therefore, taking the limit as \(n \to \infty\) for Equation (10) gives

\[
a = b \lim_{n \to \infty} P_a(X_n = b) , \quad \text{so} \quad \lim_{n \to \infty} P_a(X_n = b) = \frac{a}{b}.
\]

**Proof of claim:** Suppose that at state \(i \in \{1,2,\ldots,b-1\}\), we cannot eventually reach an absorbing state with probability 1. Let \(k\) be the state closest to 0 that we can eventually reach from state \(i\). Then from state \(k\), we cannot reach any state in \(\{0,1,\ldots,k-1\}\). Since \((X_n)\) is a martingale, \(\mathbb{E}[X_{n+1} \mid X_n = k] = k\), but since \(k\) is not an absorbing state, it means that there must be some probability of reaching a state in \(\{0,1,\ldots,k-1\}\) (otherwise, we would have \(\mathbb{E}[X_{n+1} \mid X_n = k] > k\)). Hence, we reach a contradiction. It must be that we can indeed reach absorbing state 0. By considering the highest state \(\ell < b\) that we can eventually reach from state \(i\), a similar argument can be used to prove that we can eventually reach state \(b\) from state \(i\).