STAT 150 SPRING 2010: MIDTERM EXAM

Problems by Jim Pitman. Solutions by George Chen

1. Let $X_0, Y_1, Y_2, \dots$ be independent random variables, X_0 with values in $\{0, 1, 2, \dots\}$ 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 $i = 1, 2, \dots$ For $n = 1, 2, \dots$ let

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

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

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

$$\begin{aligned} \mathbb{P}(X_{n+1} = x_{n+1} \mid X_i = x_i \text{ for } i = 0, 1, \dots, n) &= \frac{\mathbb{P}(X_i = x_i \text{ for } i = 0, 1, 2, \dots, n+1)}{\mathbb{P}(X_i = x_i \text{ for } i = 0, 1, 2, \dots, n)} \\ &= \frac{\mathbb{P}\begin{pmatrix} Y_{x_0} = 1, Y_{x_0 - 1} = \cdots = Y_{x_1 + 1} = 0, Y_{x_1} = 1, \\ Y_{x_1 - 1} = \cdots = Y_{x_2 + 1} = 0, Y_{x_2 = 1}, \dots, Y_{x_{n+1}} = 1 \end{pmatrix}}{\mathbb{P}\begin{pmatrix} Y_{x_0} = 1, Y_{x_0 - 1} = \cdots = Y_{x_1 + 1} = 0, Y_{x_1} = 1, \\ Y_{x_1 - 1} = \cdots = Y_{x_2 + 1} = 0, Y_{x_2 = 1}, \dots, Y_{x_n} = 1 \end{pmatrix}} \\ &= \frac{\mathbb{P}(Y_{x_0} = 1) \prod_{i=1}^{n+1} \left\{ \left(\prod_{j = x_i + 1}^{x_{i-1} - 1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_{x_i} = 1) \right\}}{\mathbb{P}(Y_{x_0} = 1) \prod_{i=1}^{n} \left\{ \left(\prod_{j = x_i + 1}^{x_{i-1} - 1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_{x_i} = 1) \right\}}. \end{aligned}$$

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

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

Conclude that

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_i = x_i \text{ for } i = 0, 1, \dots, n) = \frac{1}{x_n - 1}. \tag{1}$$

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

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}, X_n \le 0) = \mathbf{1}(x_{n+1} = 0). \tag{2}$$

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

$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_i = x_i \text{ for } i = 0, 1, \dots, n) = \begin{cases} \frac{1}{x_n - 1} & \text{if } x_n > 1 \text{ and } 0 < x_{n+1} < x_n, \\ \mathbf{1}(x_{n+1} = 0) & \text{if } x_n \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

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

2. For $Y_1, Y_2, \dots$ as in the previous question, let $T_0 := 0$ and for $n = 1, 2, \dots$ 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 $\{0,1,2,\dots\}$ 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 $t_{n+1} > t_n > t_{n-1} > \cdots > t_2 > 1$:

$$\begin{aligned}
&\mathbb{P}(T_{n+1} = t_{n+1} \mid T_0 = 0, T_1 = 1, T_2 = t_2, \dots, T_n = t_n) \\
&= \frac{\mathbb{P}(T_0 = 0, T_1 = 1, T_2 = t_2, \dots, T_n = t_n, T_{n+1} = t_{n+1})}{\mathbb{P}(T_0 = 0, T_1 = 1, T_2 = t_2, \dots, T_n = t_n)} \\
&= \frac{\mathbb{P}(T_0 = 0)\,\mathbb{P}(T_1 = 1 \mid T_0 = 0) \prod_{i=2}^{n+1} \left\{ \left(\prod_{j=t_{i-1}+1}^{t_i - 1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_{t_i} = 1) \right\}}{\mathbb{P}(T_0 = 0)\,\mathbb{P}(T_1 = 1 \mid T_0 = 0) \prod_{i=2}^{n} \left\{ \left(\prod_{j=t_{i-1}+1}^{t_i - 1} \mathbb{P}(Y_j = 0) \right) \mathbb{P}(Y_{t_i} = 1) \right\}}.
\end{aligned}$$

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

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

Conclude that for $n \geq 2$,

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

In particular, $\mathbb{P}(T_{n+1} = t_{n+1} \mid T_i = t_i \text{ for } i = 0,1,\dots,n)$ does not depend on $t_0, t_1, \dots, t_{n-1}$, so $\mathbb{P}(T_{n+1} \mid T_0, \dots, T_n) = \mathbb{P}(T_{n+1} \mid T_n)$, 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

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

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):

$$\mathbb{E}(Y g(X)) = \sum_x \mathbb{P}(X = x)\,\mathbb{E}(Y g(X) \mid X = x) = \sum_x \mathbb{P}(X = x)\, g(x) \underbrace{\mathbb{E}(Y \mid X = x)}_{\phi(x)} = \mathbb{E}(g(X)\,\phi(X)).$$

Next we show that ϕ is unique, i.e. if a function ϕ satisfies Equation (5), then we must have $\phi(x) = \mathbb{E}(Y \mid X = x)$. Note that the domain of ϕ 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

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(Y \mid X) \mid f(X) = f(x)) &= \mathbb{E}(\phi(X) \mid f(X) = f(x)) \\
&= \frac{\mathbb{E}(\phi(X) \mathbf{1}(f(X) = f(x)))}{\mathbb{P}(f(X) = x)} \qquad \text{(recall that } \mathbb{E}(A \mid B) = \mathbb{E}(A \mathbf{1}_B) / \mathbb{P}(B)\text{)} \\
&= \frac{\mathbb{E}(Y \mathbf{1}(f(X) = f(x)))}{\mathbb{P}(f(X) = x)} \qquad \text{(by Equation (5) where } g(x) = \mathbf{1}(f(X) = f(x))\text{)} \\
&= \mathbb{E}(Y \mid f(X) = f(x)).
\end{aligned}$$

We can verify Equation (7) with direct computation:

$$\begin{aligned}
\mathbb{E}\left(\mathbb{E}\left(Y \mid X, Z\right) \mid X = x\right) &= \sum_z \mathbb{E}\left(Y \mid X = x, Z = z\right) \mathbb{P}\left(Z = z \mid X = x\right) \\
&= \sum_z \sum_y y \mathbb{P}\left(Y = y \mid X = x, Z = z\right) \mathbb{P}\left(Z = z \mid X = x\right) \\
&= \sum_z \sum_y y \frac{\mathbb{P}\left(X = x, Y = y, Z = z\right)}{\mathbb{P}\left(X = x, Z = z\right)} \frac{\mathbb{P}\left(X = x, Z = z\right)}{\mathbb{P}\left(X = x\right)} \\
&= \sum_z \sum_y y \frac{\mathbb{P}\left(X = x, Y = y, Z = z\right)}{\mathbb{P}\left(X = x\right)} \\
&= \sum_z \sum_y y \mathbb{P}\left(Y = y, Z = z \mid X = x\right) \\
&= \sum_y y \mathbb{P}\left(Y = y \mid X = x\right) \\
&= \mathbb{E}\left(Y \mid X = x\right).
\end{aligned}$$

4. Suppose that a sequence of random variables $X_0, X_1, \ldots$ and a function f are such that

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

for every $n = 0, 1, 2, \ldots$ Explain why this implies

$$\mathbb{E}\left(f\left(X_{n+1}\right) \mid f\left(X_0\right), \ldots, f\left(X_n\right)\right) = f\left(X_n\right). \tag{9}$$

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 $\mathbf{X}^{(n)} = \begin{pmatrix} X_0 & X_1 & \cdots & X_{n-1} \end{pmatrix}^\top$ and $\mathbf{Y}^{(n)} = \begin{pmatrix} f\left(X_n\right) & 0 & \cdots & 0 \end{pmatrix}^\top$ taking on values in $\mathbb{R}^n$. Define function g by $g\left(\mathbf{X}^{(n)}\right) = \begin{pmatrix} f\left(X_0\right) & f\left(X_1\right) & \cdots & f\left(X_{n-1}\right) \end{pmatrix}^\top$. Then

$$\begin{aligned}
\mathbb{E}\left(f\left(X_n\right) \mid f\left(X_0\right), \ldots, f\left(X_{n-1}\right)\right) &= \mathbb{E}\left(\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} f\left(X_n\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| \begin{pmatrix} f\left(X_0\right) \\ f\left(X_1\right) \\ \vdots \\ f\left(X_{n-1}\right) \end{pmatrix}\right) \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mathbb{E}\left(\begin{pmatrix} f\left(X_n\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| \begin{pmatrix} f\left(X_0\right) \\ f\left(X_1\right) \\ \vdots \\ f\left(X_{n-1}\right) \end{pmatrix}\right) \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mathbb{E}\left(\mathbf{Y}^{(n)} \mid g\left(\mathbf{X}^{(n)}\right)\right) \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mathbb{E}\left(\mathbb{E}\left(\mathbf{Y}^{(n)} \mid \mathbf{X}^{(n)}\right) \mid g\left(\mathbf{X}^{(n)}\right)\right) && \text{(by Equation (6))} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mathbb{E}\left(\mathbb{E}\left(\begin{pmatrix} f\left(X_n\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| \begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_{n-1} \end{pmatrix}\right) \middle| g\left(\mathbf{X}^{(n)}\right)\right) \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \mathbb{E}\left(\begin{pmatrix} f\left(X_{n-1}\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| g\left(\mathbf{X}^{(n)}\right)\right) && \text{(by Equation (8))} \\
&= \mathbb{E}\left(\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} f\left(X_{n-1}\right) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \middle| \begin{pmatrix} f\left(X_0\right) \\ f\left(X_1\right) \\ \vdots \\ f\left(X_{n-1}\right) \end{pmatrix}\right) \\
&= \mathbb{E}\left(f\left(X_{n-1}\right) \mid f\left(X_0\right), f\left(X_1\right), \ldots, f\left(X_{n-1}\right)\right),
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}\left(f\left(X_{n+1}\right) \mid X_0, \ldots, X_n\right) &= \mathbb{E}\left(\left(\frac{q}{p}\right)^{X_{n+1}} \mid X_0, \ldots, X_n\right) \\
&= p\left(\frac{q}{p}\right)^{X_n+1} + q\left(\frac{q}{p}\right)^{X_n-1} \\
&= \frac{q^{X_n+1}}{p^{X_n}} + \frac{q^{X_n}}{p^{X_n-1}} \\
&= \frac{q^{X_n+1}}{p^{X_n}} + \frac{q^{X_n}p}{p^{X_n}} \\
&= \frac{q^{X_n}q + q^{X_n}p}{p^{X_n}} \\
&= \frac{q^{X_n}}{p^{X_n}}(q+p) \\
&= f(X_n),
\end{aligned}$$

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

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 (μ) 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 $0 \le p \le 1$. Derive the joint distribution of S and F. How can the conclusion be generalized to multinomial trials?

   Solution: Let $q = 1 - p$. We have

$$\begin{aligned}
\mathbb{P}(S = s, F = f) &= \sum_{n=0}^{\infty} \mathbb{P}(S = s, F = f \mid N = n)\,\mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S = s, N - S = f \mid N = n)\,\mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}(S = s, S = N - f \mid N = n)\,\mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} X_i = s, \sum_{i=1}^{n} X_i = n - f\right)\mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbf{1}(s = n - f)\,\mathbb{P}\left(\sum_{i=1}^{n} X_i = s\right)\mathbb{P}(N = n) \\
&= \sum_{n=0}^{\infty} \mathbf{1}(n = s + f)\,\mathbb{P}\left(\sum_{i=1}^{n} X_i = s\right)\mathbb{P}(N = n) \\
&= \mathbb{P}\left(\sum_{i=1}^{s+f} X_i = s\right)\mathbb{P}(N = s + f) \\
&= \binom{s+f}{s} p^s q^f \frac{\mu^{s+f} e^{-\mu}}{(s+f)!} \\
&= \frac{(s+f)!}{s!f!} \frac{p^s \mu^s q^f \mu^f e^{-\mu(p+q)}}{(s+f)!} \\
&= \frac{(p\mu)^s e^{-p\mu}}{s!} \frac{(q\mu)^f e^{-q\mu}}{f!} \\
&= \mathbb{P}(\text{Poisson}(p\mu) = s)\,\mathbb{P}(\text{Poisson}(q\mu) = f).
\end{aligned}$$

   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

$$\mathbb{P}\left(S_1 = s_1, S_2 = s_2, \ldots, S_k = s_k\right) = \prod_{i=1}^{k} \mathbb{P}\left(\text{Poisson}\left(p_i \mu\right) = s_i\right).$$

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

$$f_{ij} := \mathbb{P}_i\left(S_n = j \text{ for some } n \geq 1\right).$$

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 $i < j$. Deduce a formula for f_{ij} for $i = j$.

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

$$f_{ij} = \mathbb{P}_i\left(\text{hit } j \text{ before } +\infty\right) = \lim_{b\to\infty} \mathbb{P}_a\left(\text{hit } 0 \text{ before } b\right) = \left(\frac{q}{p}\right)^a = \left(\frac{q}{p}\right)^{i-j}$$

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 $f_{ij} = \mathbb{P}_i\left(\text{hit } j\right) = 1$. 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

$$f_{ij} = \lim_{b\to\infty} \mathbb{P}_a\left(\text{hit } 0 \text{ before } b\right) = \lim_{b\to\infty}\left(1 - \frac{\left(\frac{p}{q}\right)^a - 1}{\left(\frac{p}{q}\right)^b - 1}\right) = 1 - \lim_{b\to\infty} \frac{\left(\frac{p}{q}\right)^{(-i)-(-j)} - 1}{\left(\frac{p}{q}\right)^b - 1} = 1 - \lim_{b\to\infty} \frac{\left(\frac{p}{q}\right)^{-i+j} - 1}{\left(\frac{p}{q}\right)^b - 1}.$$

Since $p > q$, the right-most term's denominator goes to $+\infty$ whereas the numerator is fixed, so $\lim_{b\to\infty} \frac{\left(\frac{p}{q}\right)^{-i+j} - 1}{\left(\frac{p}{q}\right)^b - 1} = 0$. Thus, we have $f_{ij} = 1 - \lim_{b\to\infty} \frac{\left(\frac{p}{q}\right)^{-i+j} - 1}{\left(\frac{p}{q}\right)^b - 1} = 1 - 0 = 1$.

Case $i = j$: From first-step analysis, we have

$$\begin{aligned}
f_{ii} &= \mathbb{P}\left(\text{go 1 step up}\right)\mathbb{P}_{i+1}\left(\text{hit } i \text{ before } +\infty\right) + \mathbb{P}\left(\text{go 1 step down}\right)\mathbb{P}_{i-1}\left(\text{hit } i\right) \\
&= p\left(\frac{q}{p}\right)^{(i+1)-i} + q \cdot 1 \qquad \text{(using previous results)} \\
&= p\left(\frac{q}{p}\right) + q \\
&= 2q.
\end{aligned}$$

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 $P_i\left(T < \infty\right) = 1$ for all i. Derive a formula for

$$\mathbb{P}_i\left(X_{T-1} = j, X_T = k\right) \text{ for } i, j \in B^c \text{ and } k \in B$$

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:

$$\begin{aligned}
\mathbb{P}_i\left(X_{T-1} = j, X_T = k\right) &= \sum_{n=1}^{\infty} \mathbb{P}_i\left(X_{T-1} = j, X_T = k, T = n\right) \\
&= \sum_{n=1}^{\infty} P^{n-1}\left(i, j\right) P\left(j, k\right) \\
&= \left(\sum_{m=0}^{\infty} P^m\left(i, j\right)\right) P\left(j, k\right)
\end{aligned}$$

$$= \underbrace{\left(\sum_{m=0}^{\infty} Q^m(i,j)\right)}_{W(i,j)} R(j,k)$$
$$= W(i,j)\, R(j,k)\,.$$

8. In the same setting, let $f_{ij} := \mathbb{P}_i(X_n = j \text{ for some } n \geq 1)$. 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 $W_{ij} = \mathbb{E}_i(N_j)$ and $W_{jj} = \mathbb{E}_j(N_j)$. Reaching $X_n = j$ for some $n \geq 1$ is equivalent to saying that there exists a first time that we reach j; thus:

   $$f_{ij} = \mathbb{P}_i(X_n = j \text{ for some } n \geq 1) = \mathbb{P}(\text{we reach } j \text{ for the first time})\,.$$

   From first-step analysis:

   $$\mathbb{E}_i(N_j) = \mathbb{P}(\text{we reach state } j \text{ for the first time}) \cdot \mathbb{E}_j(N_j) + \mathbb{P}(\text{we never reach state } j) \cdot 0.$$

   Hence, we have

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

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 $i \in B$, $\mathbb{P}_i(T = 0) = 1$, i.e. $\boxed{\phi_i(s) = 1 \text{ for } i \in B}$. For $i \notin B$, clearly $\mathbb{P}_i(T = 0) = 0$ and for $n \geq 1$, use first-step analysis to get

   $$\begin{aligned}
   \mathbb{P}_i(T = n) &= \sum_j P(i,j)\, \mathbb{P}_j(T = n-1) \\
   &= \sum_{j \in B^c} Q(i,j)\, \mathbb{P}_j(T = n-1) + \sum_{k \in B} R(i,k)\, \mathbb{P}_k(T = n-1) \\
   &= \sum_{j \in B^c} Q(i,j)\, \mathbb{P}_j(T = n-1) + \sum_{k \in B} R(i,k)\, \mathbf{1}(n-1 = 0) \\
   &= \sum_{j \in B^c} Q(i,j)\, \mathbb{P}_j(T = n-1) + \mathbf{1}(n = 1) \sum_{k \in B} R(i,k)\,.
   \end{aligned}$$

   So

   $$\begin{aligned}
   \phi_i(s) &= \underbrace{\mathbb{P}_i(T = 0)}_{0} + \sum_{n=1}^{\infty} \mathbb{P}_i(T = n)\, s^n \\
   &= \sum_{n=1}^{\infty} \left(\sum_{j \in B^c} Q(i,j)\, \mathbb{P}_j(T = n-1) + \mathbf{1}(n = 1) \sum_{k \in B} R(i,k) \right) s^n \\
   &= \sum_{n=1}^{\infty} \sum_{j \in B^c} Q(i,j)\, \mathbb{P}_j(T = n-1)\, s^n + \sum_{n=1}^{\infty} \mathbf{1}(n = 1) \sum_{k \in B} R(i,k)\, s^n \\
   &= \sum_{j \in B^c} Q(i,j) \sum_{n=1}^{\infty} \mathbb{P}_j(T = n-1)\, s^n + \sum_{k \in B} R(i,k)\, s \\
   &= \sum_{j \in B^c} Q(i,j) \sum_{m=0}^{\infty} \mathbb{P}_j(T = m)\, s^{m+1} + \sum_{k \in B} R(i,k)\, s \\
   &= \sum_{j \in B^c} Q(i,j)\, s \sum_{m=0}^{\infty} \mathbb{P}_j(T = m)\, s^m + \sum_{k \in B} R(i,k)\, s \\
   &= s \sum_{j \in B^c} Q(i,j) \sum_{m=0}^{\infty} \mathbb{P}_j(T = m)\, s^m + s \sum_{k \in B} R(i,k)
   \end{aligned}$$

$$= s \sum_{j \in B^c} Q(i,j)\,\phi_j(s) + s \sum_{k \in B} R(i,k)$$
$$= s \sum_{j \in B^c \setminus \{i\}} Q(i,j)\,\phi_j(s) + sQ(i,i)\,\phi_i(s) + s \sum_{k \in B} R(i,k).$$

Rearranging terms gives

$$\boxed{s \sum_{j \in B^c \setminus \{i\}} Q(i,j)\,\phi_j(s) + (sQ(i,i) - 1)\,\phi_i(s) + s \sum_{k \in B} R(i,k) = 0, \qquad \text{for } i \in B^c.}$$

10. Let X be a non-negative integer valued random variable with probability generating function $\phi(s)$ for $0 \le s \le 1$. 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 ϕ and p.

Solution:

$$\begin{aligned}
\mathbb{P}(N < X) &= \sum_{x=0}^{\infty} \mathbb{P}(N < X \mid X = x)\,\mathbb{P}(X = x) \\
&= \sum_{x=0}^{\infty} \mathbb{P}(N < x)\,\mathbb{P}(X = x) \\
&= \sum_{x=0}^{\infty} \mathbb{P}(N \le x - 1)\,\mathbb{P}(X = x) \\
&= \sum_{x=0}^{\infty} \left(1 - (1-p)^x\right)\mathbb{P}(X = x) \\
&= \sum_{x=0}^{\infty} \mathbb{P}(X = x) - \sum_{x=0}^{\infty} (1-p)^x\,\mathbb{P}(X = x) \\
&= 1 - \phi(1-p).
\end{aligned}$$

11. Let X be a non-negative random variable with usual probability generating function $\phi(s)$ for $0 \le s \le 1$. Define the tail probability generating function $\tau(s)$ by

$$\tau(s) := \sum_{n=1}^{\infty} \mathbb{P}(X \ge n)\, s^n.$$

Use the identity

$$\mathbb{P}(X = n) = \mathbb{P}(X \ge n) - \mathbb{P}(X \ge n+1)$$

to derive a formula for $\tau(s)$ in terms of s and $\phi(s)$ for $0 \le s \le 1$. Discuss what happens for $s = 1$.

Solution: We have

$$\begin{aligned}
\phi(s) &= \sum_{n=0}^{\infty} \mathbb{P}(X = n)\, s^n \\
&= \sum_{n=0}^{\infty} \left(\mathbb{P}(X \ge n) - \mathbb{P}(X \ge n+1)\right) s^n \\
&= \sum_{n=0}^{\infty} \mathbb{P}(X \ge n)\, s^n - \sum_{n=0}^{\infty} \mathbb{P}(X \ge n+1)\, s^n \\
&= \mathbb{P}(X \ge 0) + \sum_{n=1}^{\infty} \mathbb{P}(X \ge n)\, s^n - \sum_{m=1}^{\infty} \mathbb{P}(X \ge m)\, s^{m-1} \\
&= \underbrace{\mathbb{P}(X \ge 0)}_{1} + \sum_{n=1}^{\infty} \mathbb{P}(X \ge n)\, s^n - s^{-1} \sum_{m=1}^{\infty} \mathbb{P}(X \ge m)\, s^m \\
&= 1 + \tau(s) - s^{-1}\tau(s) \\
&= 1 + \tau(s)\left(1 - s^{-1}\right),
\end{aligned}$$

so

$$\tau(s) = \frac{\phi(s) - 1}{1 - s^{-1}}.$$

It is clear that by the definition of $\tau(s)$, when $s = 1$, we have $\tau(1) = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) = \mathbb{E}(X)$. We can also see this via l'Hopital's rule:

$$\lim_{s \to 1} \tau(s) = \lim_{s \to 1} \frac{\phi(s) - 1}{1 - s^{-1}} = \lim_{s \to 1} \frac{\phi'(s)}{s^{-2}} = \frac{\phi'(1)}{1} = \phi'(1) = \mathbb{E}(X).$$

12. Consider a random walk on the 3 vertices of a triangle labeled clockwise 0, 1, 2. At each step, the walk moves clockwise with probability p and counter-clockwise with probability q, where $p + q = 1$. Let P denote the transition matrix. Observe that

$$P^2(0,0) = 2pq; \qquad P^3(0,0) = p^3 + q^3; \qquad P^4(0,0) = 6p^2q^2.$$

Derive a similar formula for $P^5(0,0)$.

Solution: Consider a $p \uparrow, q \downarrow$ random walk on $\mathbb{Z}$. Modulo 3, we are traversing the triangle described. We restrict the rest of our discussion to the random walk on $\mathbb{Z}$ where we start at the origin and want to reach state 0 of the triangle (i.e. any multiple of 3 for the random walk on $\mathbb{Z}$) in 5 steps. Observe that in 5 steps, we cannot possibly reach any multiple of 3 larger than 3 away from the origin. Also, since we move an odd number of steps, we cannot return to the origin. However, we can reach $+3$ (4 up and 1 down in any combination) and -3 (4 down and 1 up in any combination). Therefore,

$$P^5(0,0) = \underbrace{\binom{5}{1}}_{\substack{\text{in 5 moves,} \\ \text{1 is down and} \\ \text{the rest are up}}} p^4 q + \underbrace{\binom{5}{1}}_{\substack{\text{in 5 moves,} \\ \text{1 is up and the} \\ \text{rest are down}}} pq^4 = 5p^4q + 5pq^4.$$

13. A branching process with Poisson (λ) offspring distribution started with one individual has extinction probability p with $0 < p < 1$. Find a formula for λ 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

$$\lim_{n \to \infty} \mathbb{P}_a(X_n = b)$$

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

$$a = \mathbb{E}[X_n] = \sum_{i=0}^{b} i\mathbb{P}_a(X_n = i) = \sum_{i=1}^{b-1} i\mathbb{P}_a(X_n = i) + b\mathbb{P}_a(X_n = b). \tag{10}$$

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 $i \in \{1, 2, \ldots, b-1\}$, $\lim_{n\to\infty} \mathbb{P}_a(X_n = i) = 0$. Therefore, taking the limit as $n \to \infty$ for Equation (10) gives

$$a = b \lim_{n \to \infty} \mathbb{P}_a(X_n = b), \quad \text{so} \quad \lim_{n \to \infty} \mathbb{P}_a(X_n = b) = \frac{a}{b}.$$

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.