1. Prove that for random variables X and Y, $\mathbb{E}[X] = \sum_y \mathbb{E}[X \mid Y = y] \cdot p(Y = y)$ (where the summation is over all values y that the random variable Y can take).

Solution:

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_x x \cdot p(X = x) \\
&= \sum_x x \sum_y p(X = x \mid Y = y) \cdot p(Y = y) \\
&= \sum_y \sum_x x \cdot p(X = x \mid Y = y) \cdot p(Y = y) \\
&= \sum_y \mathbb{E}[X|Y = y].
\end{aligned}
$$

2. Prove the "law of iterated expectation", that is, for random variables X and Y, $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X, Y]]$.

Solution:

$$
\begin{aligned}
\mathbb{E}[\mathbb{E}[X \mid Y]] &= \mathbb{E}\left[\sum_x x \cdot p(X = x \mid Y = y) \cdot p(Y = y)\right] \\
&= \mathbb{E}\left[\sum_x x \cdot \frac{p(X = x, Y = y)}{P(Y = y)} \cdot p(Y = y)\right] \\
&= \mathbb{E}\left[\sum_x x \cdot p(X = x)\right] \\
&= \mathbb{E}[\mathbb{E}[X]] \\
&= \mathbb{E}[X].
\end{aligned}
$$

3. Prove that $\mathrm{Var}[X] := \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (E[X])^2$.

Solution:

$$\begin{aligned}\mathbb{E}[(X - \mathbb{E}[X])^2] &= \mathbb{E}[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X]^2)] \\ &= \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[x] + +(\mathbb{E}[X]^2) \\ &= \mathbb{E}[X^2] - (E[X])^2.\end{aligned}$$

4. Prove that $\mathrm{Cov}[X, Y] = 0$ when X and Y are independent.

Solution: As X and Y are independent, $E[XY] = E[X]E[Y]$. Thus the answer is clear.

5. Let X be a random variable defined as $X = \sum_{i=1}^{n} X_i$ where each X_i is a Bernoulli random variable, and all X_i are independent. Let $p_i = p(X_i = 1)$. Prove that for any $\delta \geq e^2 - 1$,

$$p(X \geq (1+\delta)\mathbb{E}[X]) \leq e^{-(\delta+2)\mathbb{E}[X]}.$$

Solution: By Chernoff's inequality, we have for any $\delta > 0$

$$p(X \geq (1+\delta)\mathbb{E}[X]) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mathbb{E}[X]}.$$

As

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \geq 1$$

and $E[X] \geq 0$, we move to upper bound

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}.$$

Let $\delta \geq e^2 - 1$, then $1 + \delta \geq e^2$ and so we have

$$\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \leq \frac{e^{\delta}}{e^{2(1+\delta)}} = e^{-2(\delta+1)},$$

giving the desired result.

6. Consider the following probabilistic process. You have a collection of n many "bins", and also n many "balls". Assume that both are numbered $1, 2, \ldots, n$. The balls are thrown into the bins independently and uniformly at random, that is, for each ball, every bin has the same probability of being chosen for the ball.

   Formally prove that with high probability (of the form $1 - 1/n^c$ for some constant $c \in \mathbb{R}_{\geq 1}$) the maximum number of balls in any bin is $O(\log n)$.

> **Solution:**

7. Let $X_1, X_2, \ldots$ be random variables such that $Z_1, Z_2, \ldots$ with $Z_n = \sum_{i=1}^{n} X_i$ are a martingale. Show that $\mathbb{E}[X_i \cdot X_j] = 0$ whenever $i \neq j$.

> **Solution:** Without loss of generality, we assume $j < i$. As $X_i = Z_i - Z_{i-1}$, we have
> $$\mathbb{E}[X_i X_j] = \mathbb{E}[(Z_i - Z_{i-1})X_j].$$
> By the law of iterated expectation, we have
> $$\begin{aligned}\mathbb{E}[(Z_i - Z_{i-1})X_j] &= \mathbb{E}[\mathbb{E}[(Z_i - Z_{i-1})X_j \mid Z_1, \ldots, Z_n]] \\ &= \mathbb{E}[X_j \mathbb{E}[Z_i - Z_{i-1} \mid Z_1, \ldots, Z_n]] \\ &= \mathbb{E}[X_j(0)] = 0.\end{aligned}$$

8. Let $X_0, X_1, X_2, \ldots$ be a sequence of random variables with finite expectations, satisfying
   $$\mathbb{E}[X_{n+1} \mid X_0, X_1, \ldots, X_n] = aX_n + bX_{n-1}$$
   for $n \geq 1$, where $0 < a, b < 1$ and $a + b = 1$. Find a value of β for which $Z_n = \beta X_n + X_{n-1}$ ($n \geq 1$) defines a martingale with respect to $X_0, X_1, X_2, \ldots$.

Solution: Each X_i has finite expectation, so each Z_i does too. Assume that such a $\beta \in \mathbb{R}$ exists. Then

$$\begin{aligned}
\mathbb{E}[Z_{n+1} \mid X_0, \ldots, X_n] &= \mathbb{E}[\beta X_{n+1} + X_n \mid X_0, \ldots, X_n] \\
&= \mathbb{E}[\beta X_{n+1} \mid X_0, \ldots, X_n] + X_n \\
&= \beta(aX_n + bX_{n-1}) + X_n \\
&= (a\beta + 1)X_n + b\beta X_{n-1} \\
&= \beta X_n + X_{n-1}.
\end{aligned}$$

Giving us $\beta = \frac{1}{\beta}$. This can be checked to satisfy the requirements of a martingale (note that clearly Z_n is a function of $\{X_i\}_{i=0}^{n}$).

9. Consider a bag that at time $n = 0$ contains one red marble and one blue marble. At each time step $n \geq 1$ you inspect the colour a randomly chosen marble and add an extra marble of that colour to the bag. Let M_n denote the fraction of red marbles after time step n. Is $(M_n)_{n\geq 0}$ a martingale? Prove your answer.

Solution: We let S_n be the number of red marbles in the bag after step n. We see that $M_n = \frac{S_n}{n+2}$, thus we can swap the conditioning on one with the conditioning with the other.

$$\begin{aligned}
\mathbb{E}[M_{n+1} \mid M_1, \ldots M_n] &= \mathbb{E}\left[\frac{S_{n+1}}{n+3} \,\middle|\, M_1, \ldots M_n\right] \\
&= \frac{1}{n+3}\mathbb{E}\left[S_{n+1} \mid M_1, \ldots M_n\right] \\
&= \frac{1}{n+3}\mathbb{E}\left[S_{n+1} \mid S_1, \ldots S_n\right].
\end{aligned}$$

We note that

$$\begin{aligned}
\mathbb{E}\left[S_{n+1} \mid S_n\right] &= \Pr[S_{n+1} = S_n + 1](S_n + 1) + \Pr[S_{n+1} = S_n](S_n) \\
&= \frac{S_n}{n+2}(S_n + 1) + \left(1 - \frac{S_n}{n+2}\right) S_n \\
&= \frac{S_n}{n+2}\left(S_n + 1 + n + 2 - S_n\right) \\
&= \frac{S_n(n+3)}{n+2}.
\end{aligned}$$

Thus

$$\mathbb{E}[M_{n+1} \mid M_1, \dots M_n] = \frac{1}{n+3} \cdot \frac{S_n(n+3)}{n+2} = M_n,$$

as required. We have expressed M_n as a function of S_n, and M_n is bounded for all n. Thus $(M_n)_{n\in\mathbb{N}}$ is a martingale.

10. The game "Next Is Black" can be played by a single player with a shuffled but otherwise ordinary pack of 52 cards. At times $n = 1, 2, \dots, 52$ the player turns over a new card and notes its colour. Just once in the game they must say, just before turning over a card, "Next Is Black". They win the game if that card is indeed black. Let B_n be the number of black cards remaining face down after the n-th card has been exposed. Show that $X_n = B_n/(52-n)$, $0 \leq n \leq 52$, defines a martingale. Show that there is no strategy for the player which would result in a winning probability different from $1/2$.

Solution: First, we observe that $\mathbb{E}[B_{n+1} \mid B_n] = B_n - \frac{B_n}{(52-n)}$. Then

$$\begin{aligned}
\mathbb{E}[X_{n+1} \mid B_0, \ldots, B_n] &= \mathbb{E}\left[\frac{B_{n+1}}{52-(n+1)} \,\middle|\, B_0, \ldots, B_n\right] \\
&= \frac{1}{52-(n+1)} \mathbb{E}\left[B_{n+1} \mid B_0, \ldots, B_n\right] \\
&= \frac{1}{52-(n+1)} \left(B_n - \frac{B_n}{(52-n)}\right) \\
&= \frac{1}{52-(n+1)} \left(\frac{(52-(n+1))B_n}{(52-n)}\right) \\
&= X_n.
\end{aligned}$$

For all $n \in \mathbb{N}_0$, X_n is clearly bounded (finite game), and we have expressed it as a function of $(B_i)_{i=0}^n$. Thus X_n is a martingale. A strategy in this game corresponds to a stopping time, so let T be a stopping time. As this is a finite game, T must be bounded. Thus by the optional stopping theorem

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] = \frac{1}{2}.$$

We note that the probability of winning is given by $\mathbb{E}[X_T]$, and thus any strategy has probability of winning $1/2$.

11. Consider a time-homogeneous Markov chain $(X_0, X_1, \ldots)$ with initial distribution $\boldsymbol{x}^{(0)} = (1, 0, 0, 0)$ and transition matrix

$$P = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}.$$

(a) Calculate P^2. How can P^2 (in general) be interpreted?

Solution:

$$P^2 = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

$P^2[i,j]$ denotes the two-step transition probability; that is,

$$\Pr[X_{t+2} = j \mid P_t = i] = P^2[i,j].$$

(b) Prove that

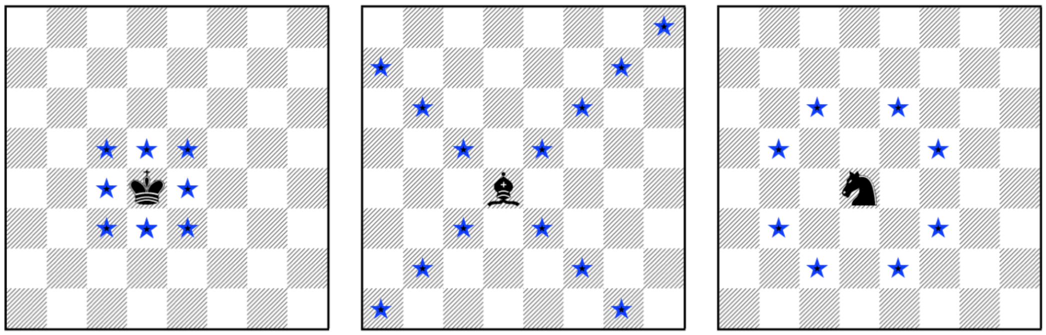
$$\boldsymbol{x}^{(n)} = \begin{cases} \left(0, \frac{1}{2}, 0, \frac{1}{2}\right) & \text{for } n = 1, 3, 5, \ldots \\ \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right) & \text{for } n = 2, 4, 6, \ldots \end{cases}$$

Solution: We have $\boldsymbol{x}^{(t+k)} = \boldsymbol{x}^{(t)} P^k$ for all $k \in \mathbb{N}$. Let $P = P'$. Note that $P^2 = P'$ and $P'P = P$. Thus $P^{2n} = P'$ and $P^{2n+1} = P$ for all $n \in \mathbb{N}$. Thus

$$\begin{aligned} \boldsymbol{x}^{(2n)} &= \boldsymbol{x}^{(0)} P^{2n} \\ &= \boldsymbol{x}^{(0)} P' \\ &= \left(\tfrac{1}{2}, 0, \tfrac{1}{2}, 0\right) \\ \boldsymbol{x}^{(2n+1)} &= \boldsymbol{x}^{(0)} P^{2n+1} \\ &= \boldsymbol{x}^{(0)} P \\ &= \left(0, \tfrac{1}{2}, 0, \tfrac{1}{2}\right). \end{aligned}$$

12. We model the movement of a single chess figure on a chess board as a (time-homogeneous) Markov chain. Let the state space be defined by the set of squares $S = \{s_1, \ldots, s_{64}\}$. Let X_n denote the position of the piece at time n. Let the transition matrix be defined by uniformly choosing from all possible next steps. (You do

not need to explicitly provide P.)



Determine if the Markov chain $(X_0, X_1, \ldots)$ is irreducible and/or aperiodic if the chess piece in question is

(a) a king;

> **Solution:** It is clear that a king can move to any square, thus the Markov chain is irreducible. The king can move twice and be back at the original square, and similarly it can move thrice and be back at the origin square. Furthermore, the king can do this at any square, thus it is aperiodic.

(b) a bishop; or

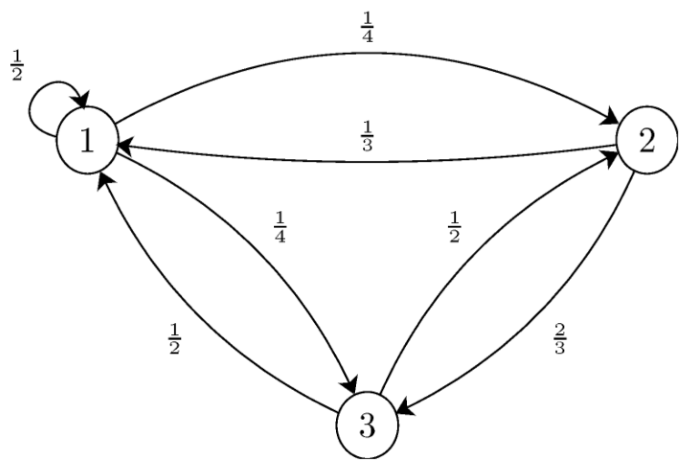
> **Solution:** A bishop can only move to the colour it started on, thus it is reducible. Using a similar to reasoning to before, it is aperiodic.

(c) a knight.

> **Solution:** A knight may move to any square, thus it is irreducible. A knight can only move to a different colour than the colour it is on, thus the period of any state must be a multiple of 2. But a knight can move to a piece and return back regardless on the square it is on. Thus it is periodic with period 2.

The stars in the figure above depict the possible single-move destinations for each of the three pieces.

13. Consider the Markov chain shown below.



(a) Is this chain irreducible?

> **Solution:** Yes, every state is reachable regardless of which state you are on.

(b) Is this chain aperiodic?

> **Solution:** Yes, we do a two-step transition $2 \to 3 \to 2$ and a three-step transition $2 \to 1 \to 3 \to 2$, so the period of 2 must be 1. For 3, we can do a two-step transition $3 \to 2 \to 3$ and a three-step transition $3 \to 1 \to 2 \to 3$, sot he period of 3 must be 1. Finally for 1, we have the one-step transition $1 \to 1$. All of these transitions have non-zero probability.

(c) Find the stationary distribution for this chain.

> **Solution:** We construct the transition matrix P for this chain:
>
> $$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$
>
> We are looking for a left eigenvector $\boldsymbol{\pi}$ of P with eigenvalue 1 (which corresponds to a right eigenvector of $P^\intercal$ with eigenvalue 1) such that $\|\boldsymbol{\pi}\|_1 = 1$. Let $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3)$ such that $\boldsymbol{\pi} P = \pi$. Then
>
> $$\left(\tfrac{1}{2}\pi_1 + \tfrac{1}{3}\pi_2 + \tfrac{1}{2}\pi_3, \tfrac{1}{4}\pi_1 + \tfrac{1}{2}\pi_3, \tfrac{1}{4}\pi_1 + \tfrac{2}{3}\pi_2\right) = (\pi_1, \pi_2, \pi_3).$$

That is, the system of equations

$$\begin{aligned}
-\tfrac{1}{2}\pi_1 + \tfrac{1}{3}\pi_2 + \tfrac{1}{2}\pi_3 &= 0,\\
\tfrac{1}{4}\pi_1 + -\pi_2 + \tfrac{1}{2}\pi_3 &= 0,\\
\tfrac{1}{4}\pi_1 + \tfrac{2}{3}\pi_2 - \pi_3 &= 0,\\
\pi_1 + \pi_2 + \pi_3 &= 1.
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
3\pi_1 - 2\pi_2 - 3\pi_3 &= 0,\\
\pi_1 + -4\pi_2 + 2\pi_3 &= 0,\\
3\pi_1 + 8\pi_2 - 12\pi_3 &= 0,\\
\pi_1 + \pi_2 + \pi_3 &= 1.
\end{aligned}$$

Putting this into reduced row echelon form, we get

$$\left(\begin{array}{ccc|c} 3 & -2 & -3 & 0\\ 1 & -4 & 2 & 0\\ 3 & 8 & -12 & 0\\ 1 & 1 & 1 & 1 \end{array}\right) = \left(\begin{array}{ccc|c} 35 & 0 & 0 & 16\\ 0 & 35 & 0 & 9\\ 0 & 0 & 7 & 2\\ 0 & 0 & 0 & 0 \end{array}\right).$$

Thus we have

$$\pi_1 = \tfrac{16}{35}, \qquad \pi_2 = \tfrac{9}{35}, \qquad \pi_3 = \tfrac{2}{7}.$$

14. Consider a state space $S = \{s_1, \dots, s_m\}$ and two distributions $\boldsymbol{\sigma}, \boldsymbol{\tau}$ on S. Recall that the total variation distance $d_{\mathrm{TV}}(\boldsymbol{\sigma}, \boldsymbol{\tau})$ is define to be $d_{\mathrm{TV}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \frac{1}{2}\|\boldsymbol{\sigma} - \boldsymbol{\tau}\|_1$. Prove that

$$d_{\mathrm{TV}}(\boldsymbol{\sigma}, \boldsymbol{\tau}) = \max_{A \subset S} |\sigma(A) - \tau(A)|$$

where $\sigma(A) = \sum_{i \in A} \sigma_i$ and $\tau(A) = \sum_{i \in A} \tau_i$.

Solution: