1. Reversibility and the Cut Property

a. A cut of a graph is a partition of its states $S$ into two disjoint subsets $(T, S \setminus T)$. Show that for an irreducible Markov chain at stationarity, flow-in equals flow-out holds across any cut of the Markov chain.

Hint: try induction on the size of one of the subsets of the cut, and write out the flow equations at each step.

b. We can convert the transition diagram of any Markov chain, which is a directed graph, into an undirected graph as follows. Every directed edge is replaced with an undirected edge; if $i$ has a directed edge to $j$ and vice versa, there is still only one undirected edge between them.

Assuming the Markov chain is irreducible and positive recurrent, a sufficient condition for the detailed balance equations to hold is that the resulting graph is a tree. Explain why this is true in words.

Solution:

a. For simplicity, because every $T \subseteq S$ uniquely determines $(T, S \setminus T)$, we will refer to just $T$ as a cut. For all cuts of size 1, the claim is true by the global balance equations. Now assume that the flow equations hold for all cuts of size $k$. Observe that it is possible to construct all cuts of size $k + 1$ by adding the appropriate node to a cut of size $k$, so take without loss of generality the cut $\{1, \ldots, k\}$ and the new state $k + 1$. Let $F_{in}([k])$ be the flow into the cut $\{1, \ldots, k\}$ and $F_{out}([k])$ the flow out. Then

$$F_{in}([k + 1]) = F_{in}([k]) - \sum_{j=1}^{k} \pi(k + 1) \cdot p(k + 1, j) + \sum_{j=k+2}^{n} \pi(j) \cdot p(j, k + 1)$$

$$F_{out}([k + 1]) = F_{out}([k]) - \sum_{j=1}^{k} \pi(j) \cdot p(j, k + 1) + \sum_{j=k+2}^{n} \pi(k + 1) \cdot p(k + 1, j).$$

By the inductive hypothesis, $F_{in}([k]) = F_{out}([k])$. Moreover, the remaining terms exactly match the balance equations of state $k + 1$. Thus we have $F_{in}([k + 1]) - F_{out}([k + 1]) = 0$, and the induction is complete.

Alternate solution. The solution above technically only holds for cuts where one subset is finite. However, more generally, the cut property holds for any cut $(T, S \setminus T)$ of any (possibly reducible) chain at stationarity:

$$\sum_{i \in T} \sum_{j \in S \setminus T} \pi(i) \cdot p(i, j) = \sum_{i \in T} \sum_{j \in S \setminus T} \pi(j) \cdot p(j, i).$$
We observe that stationarity, or global balance, gives the equality

$$
\sum_{i \in T} \left( \sum_{j \in S} \pi(i) \cdot p(i, j) \right) = \sum_{i \in T} \pi(i) = \sum_{i \in T} \left( \sum_{j \in S} \pi(j) \cdot p(j, i) \right).
$$

Intuitively, this says that the total mass in all states in $T$ remains constant; the net flows into $T$ and out of $T$ cancel out. But now, subtracting the flow that remains entirely within $T$ from both sides of the equality, we are done. This quantity is

$$
\sum_{i \in T} \sum_{j \in T} \pi(i) \cdot p(i, j) = \sum_{i \in T} \sum_{j \in T} \pi(j) \cdot p(j, i),
$$

and we see that what remains is the flow from $T$ to $S \setminus T$ equalling the flow to $T$ from $S \setminus T$, as desired. Note that we did not need to use irreducibility anywhere. Informally, we have simply shown that $\text{flow}(T, S) = \text{flow}(S, T)$ from global balance implies

$$
\text{flow}(T, S) - \text{flow}(T, T) = \text{flow}(T, S \setminus T) = \text{flow}(S \setminus T, T) = \text{flow}(S, T) - \text{flow}(T, T).
$$

b. The previous part tells us if a Markov chain is at stationarity, then the net flow across every cut must be 0. Trees have the unique property that for every edge, there exists a cut through only that edge. Thus, if the chain’s undirected graph is a tree, every pair of states has a cut that only goes through them. In this case, the net flow property becomes precisely the detailed balance equations.
2. Poisson Process Warmup

Give an interpretation of the following fact in terms of a Poisson process with rate $\lambda$. If $N$ is Geometric with parameter $p$ and $(X_k)_{k \in \mathbb{N}}$ are i.i.d. Exponential($\lambda$), then $X_1 + \cdots + X_N$ has an Exponential distribution with parameter $\lambda p$.

**Solution:** Consider a Poisson process with rate $\lambda$, and split the process by keeping each arrival independently with probability $p$. In the original process, the interarrival times are i.i.d. Exponential($\lambda$), and $X_1 + \cdots + X_N$ represents the amount of time until the first arrival we keep. By Poisson splitting, we know that the split process is a Poisson process with rate $\lambda p$, so the time until its first arrival is an Exponential random variable with parameter $\lambda p$. 
3. Poisson Process Arrival Times

Consider a Poisson process \((N_t)_{t \geq 0}\) with rate \(\lambda = 1\). For \(i \in \mathbb{Z}^+\), let \(T_i\) be the time of the \(i\)th arrival.

a. Find \(\mathbb{E}(T_3 \mid N(1) = 2)\).

b. Given \(T_3 = s\), where \(s > 0\), find the joint distribution of \(T_1\) and \(T_2\).

c. Find \(\mathbb{E}(T_2 \mid T_3 = s)\).

Solution:

a. By the memoryless property, \(\mathbb{E}(T_3 \mid N(1) = 2) = 1 + \mathbb{E}(T_1) = 1 + \lambda^{-1} = 2\).

b. As the distribution of the sum of i.i.d. Exponential random variables is Erlang,

\[
f_{T_i}(s) = \frac{s^{i-1}e^{-s}}{(i-1)!} \mathbb{1}_{s \geq 0}.
\]

Then, by Bayes’ rule and the memorylessness of Exponential distributions,

\[
f_{T_1, T_2 \mid T_3}(s_1, s_2 \mid s) = \frac{f_{T_1, T_2, T_3}(s_1, s_2, s)}{f_{T_3}(s)}
= \frac{e^{-s_1}e^{-(s_2-s_1)}e^{-(s-s_2)}}{s^2e^{-s}/2!} \mathbb{1}_{\{0 \leq s_1 \leq s_2 \leq s\}}
= \frac{2}{s^2} \mathbb{1}_{\{0 \leq s_1 \leq s_2 \leq s\}}.
\]

In other words, \(T_1\) and \(T_2\) are uniformly distributed on the feasible region \(\{0 \leq s_1 \leq s_2 \leq s\}\). In particular, the joint distribution is precisely that of the order statistics of 2 i.i.d. Uniform([0, s]) random variables.

c. By part b, \(T_2\) is the maximum of 2 Uniform([0, s]) random variables. Thus, for \(0 \leq x \leq s\),

\[
F_{T_2 \mid T_3}(x \mid s) = \mathbb{P}(T_2 \leq x \mid T_3 = s) = \left(\frac{x}{s}\right)^2
\]

\[
f_{T_2 \mid T_3}(x \mid s) = \frac{2x}{s^2} \mathbb{1}_{0 \leq x \leq s}
\]

\[
\mathbb{E}(T_2 \mid T_3 = s) = \int_0^s \frac{2x^2}{s^2} \, dx = \frac{2s}{3}.
\]
4. **System Shocks**

For a positive integer $n$, let $X_1, \ldots, X_n$ be independent Exponentially distributed random variables, each with mean 1. Let $\gamma > 0$. A system experiences shocks at times $k = 1, \ldots, n$, and the size of the shock at time $k$ is $X_k$.

a. Suppose that the system fails if any shock exceeds the value $\gamma$. What is the probability of system failure?

b. Suppose instead that the effect of the shocks is cumulative, i.e. the system fails when the total amount of shock received exceeds $\gamma$. What is the probability of system failure?

**Solution:**

a. The system fails if $\max\{X_1, \ldots, X_n\} > \gamma$, so

$$
P(\max\{X_1, \ldots, X_n\} > \gamma) = 1 - P(\max\{X_1, \ldots, X_n\} \leq \gamma)
= 1 - \prod_{k=1}^{n} P(X_k \leq \gamma) = 1 - (1 - e^{-\gamma})^n.
$$

b. $P(X_1 + \cdots + X_n > \gamma) = P(N_\gamma < n)$, where $(N_t)_{t \geq 0}$ is a Poisson process with rate 1, so

$$
P(X_1 + \cdots + X_n > \gamma) = \sum_{k=0}^{n-1} \frac{\gamma^k}{k!} e^{-\gamma}.
$$
5. Sum-Quota Sampling

Consider the problem of estimating the mean interarrival time of a Poisson process \((N_t)_{t \geq 0}\) with rate \(\lambda\), where \(N_t\) denotes the number of arrivals by time \(t\). In sum-quota sampling, the number of samples is not fixed in advance; instead, we wait until a fixed time \(t\), and take the average of the interarrival times seen so far. If we let \(\tau_i\) denote the \(i\)th interarrival time, then

\[
\bar{\tau} := \frac{\tau_1 + \cdots + \tau_{N_t}}{N_t}.
\]

Of course, the above quantity is not defined when \(N_t = 0\), so we must condition on the event \(\{N_t > 0\}\). Compute \(E(\bar{\tau} \mid N_t > 0)\).

Solution: We proceed by conditioning on the values of \(N_t\). Note that for \(n \in \mathbb{Z}^+\),

\[
P(N_t = n \mid N_t > 0) = \frac{1}{1 - e^{-\lambda t}} \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
\]

Now, by the law of total expectation,

\[
E(\bar{\tau} \mid N_t > 0) = \sum_{n=1}^{\infty} E \left( \frac{\tau_1 + \cdots + \tau_{N_t}}{N_t} \mid N_t = n \right) \cdot P(N_t = n \mid N_t > 0)
\]

Recall that conditioned on the number of arrivals in a given interval, the arrival times in the interval have the same joint distribution as i.i.d. uniform order statistics on the interval. Thus, conditioned on \(\{N_t = n\}\), the sum \(\tau_1 + \cdots + \tau_{N_t}\) is the maximum of \(n\) Uniform([0, \(t\)]) random variables, which has CDF \(\left(\frac{x}{t}\right)^n\), PDF \(\left(\frac{n}{t}\right)\left(\frac{x}{t}\right)^{n-1}\) for \(x \in [0, t]\), and expectation \(tn^{n+1}\).

\[
= \sum_{n=1}^{\infty} \frac{1}{n} \frac{tn}{n+1} \cdot \frac{1}{1 - e^{-\lambda t}} \frac{(\lambda t)^n}{n!} e^{-\lambda t}
\]

\[
= \frac{e^{-\lambda t}}{\lambda (1 - e^{-\lambda t})} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n+1}}{(n+1)!}
\]

\[
= \frac{e^{-\lambda t}}{\lambda (1 - e^{-\lambda t})} (e^{\lambda t} - 1 - \lambda t)
\]

\[
= \frac{1}{\lambda} \left( 1 - \frac{\lambda t e^{-\lambda t}}{1 - e^{-\lambda t}} \right).
\]

The expectation \(E(\bar{\tau} \mid N_t > 0)\) does not quite equal \(\frac{1}{\lambda}\), the true mean that we want to estimate. However, it will indeed tend to \(\frac{1}{\lambda}\) as \(t\) increases.
6. Basketball II

Captain America and Superman are playing an untimed basketball game in which the two players score points according to independent Poisson processes with rates $\lambda_C$ and $\lambda_S$ respectively. The game is over when one player has scored $k$ more points than the other.

a. Suppose $\lambda_C = \lambda_S$, and suppose Captain America has a head start of $m < k$ points. Find the probability that Captain America wins.

Hint: if $\alpha_i = \frac{1}{2} \alpha_{i-1} + \frac{1}{2} \alpha_{i+1}$, then $\alpha_{i+1} - \alpha_i = \alpha_i - \alpha_{i-1}$.

b. Keeping the assumptions, find the expected time $E(T)$ it will take for the game to end.

Hint: consider the telescoping sum $\beta_j = \beta_0 + (\beta_1 - \beta_0) + \cdots + (\beta_j - \beta_{j-1})$.

Solution:

a. Consider the merged process with rate $\lambda_C + \lambda_S$. We see that each point is one for Captain America with probability $p := \frac{\lambda_C}{\lambda_C + \lambda_S}$ and one for Superman with probability $1 - p$. Then, the Markov chain whose state is the number of additional points Superman needs to score to win has transition probabilities

$$
p(0,0) = 1,
p(i,i + 1) = p, \text{ where } 0 < i < 2k,
p(i,i - 1) = 1 - p, \text{ where } 0 < i < 2k,
p(2k,2k) = 1.
$$

As $\lambda_C = \lambda_S$, i.e. $p = \frac{1}{2}$, this is also known as the symmetric gambler’s ruin problem for $n = 2k$, which has the following transition diagram:

![Transition Diagram](attachment:image.png)

We are interested in the hitting probabilities $\alpha_i$, the probability of eventually reaching the absorbing state $2k$ starting from $i$, which we can find by first-step analysis. Writing $q = 1 - p$, we have the system of equations

$$
\alpha_i = p \cdot \alpha_{i+1} + q \cdot \alpha_{i-1} = \frac{1}{2} \alpha_{i-1} + \frac{1}{2} \alpha_{i+1}
$$

with boundary conditions $\alpha_0 = 0$ and $\alpha_{2k} = 1$. Now, drawing out the values $\alpha_0, \ldots, \alpha_{2k}$ on the number line $[0, 1]$, we find that they are in fact evenly spaced, with each $\alpha_i$ being the midpoint of $[\alpha_{i-1}, \alpha_{i+1}]$. Thus $\alpha_i$ is directly proportional to $i$, the “distance” of state $i$ from 0, and we find the final answer of

$$
P(Captain America wins) = \alpha_{m+k} = \frac{m + k}{2k}.
$$

Remark: the symmetric case of $p = \frac{1}{2}$ is in fact a special case. For $p \neq \frac{1}{2}$, we define the ratio $r = \frac{q}{p}$ and use the recurrence $\alpha_{i+1} - \alpha_i = r(\alpha_i - \alpha_{i-1})$, which gives a geometric sum in terms of $\alpha_1$ by telescoping. This method fails precisely for $r = 1$, or $p = q = \frac{1}{2}$. 

7
b. In the continuous-time Markov chain above, note that the holding time \( \tau_n \) for each jump is i.i.d. as Exponential\((2\lambda)\), where \( \lambda = \lambda_C = \lambda_S \). If \( N_i \) is the number of jumps made until the game ends, starting from state \( i \), then Wald’s identity, or the law of total expectation with independence, tells us that

\[
\mathbb{E}(T) = \mathbb{E} \left( \sum_{n=1}^{N_i} \tau_n \right) = \mathbb{E}(N_j \cdot \mathbb{E}(\tau_1)) = \mathbb{E}(N_j) \cdot \mathbb{E}(\tau_1) = \frac{\mathbb{E}(N_j)}{2\lambda}.
\]

To compute \( \beta_i := \mathbb{E}(N_i) \), let \( \Delta_i := \mathbb{E}(N_{i+1}) - \mathbb{E}(N_i) \). The first-step equations are

\[
\beta_i = 1 + \frac{1}{2} \beta_{i-1} + \frac{1}{2} \beta_{i+1}
\]

with \( \beta_0 = \beta_2 = 0 \), which we can rewrite in terms of the difference as \( \Delta_i = \Delta_{i-1} - 2 \). In particular, we have \( \Delta_{2k-1} = \Delta_0 - 2(2k-1) \), and therefore

\[
-\beta_{2k-1} = \Delta_{2k-1} = \Delta_0 - 2(2k-1) = \beta_1 - 2(2k-1).
\]

But \( \beta_{2k-1} = \beta_1 \) by symmetry, so \( \beta_1 = 2k-1 = \beta_1 - \beta_0 = \Delta_0 \), and the previous recurrence gives \( \Delta_i = \Delta_0 - 2i = 2k - 1 - 2i \). To calculate \( \beta_i \), we use telescopic cancellation:

\[
\beta_j = \beta_0 + \sum_{i=0}^{j-1} (\beta_{i+1} - \beta_i) = \sum_{i=0}^{j-1} \Delta_i = \sum_{i=0}^{j-1} (2k - 1 - 2i) = j(2k - j).
\]

As \( j = m + k \) was our starting state, we have \( \mathbb{E}(N_{m+k}) = (k + m)(k - m) \), and thus

\[
\mathbb{E}(T) = \frac{(k + m)(k - m)}{2\lambda}.
\]