# *Probability and Computation*: Problem Sheet 4 Solutions

**You are encouraged to submit your solutions at student reception or by emailing them to jas289 by 2pm Friday 21th of February**

**Question 1.** *Let $X_n$ be the sum of $n$ independent rolls of a fair die. Show that, for any $k \geq 2$,*

$$\lim_{n\to\infty} \mathbf{P}[\, X_n \text{ is divisible by } k\,] = \frac{1}{k}.$$

**Question 2.** *When the U bus arrives outside the Computer Lab, the next bus arrives in $1, 2, \ldots, 20$ minutes with equal probability. You arrive at the bus stop without checking the schedule, at some fixed time $n$.*

*(i)* *How could you model $X_n$, the number of minutes until the next bus when you arrive at time $n$, as a Markov chain?*

*(ii)* *Buses have been coming and going all day so we can assume the chain has mixed when you arrive. What is the probability of waiting $i$ minutes for a bus in relation to the Chain?*

*(iii)* *How long, on average, do you wait until the next bus arrives?*

*(iv)* *What is the standard deviation of this time?*

---

*Solution: Item (i):* Let $X_n$ be the time elapsed from time $n$ till the arrival of the next bus. Then $X_n$ is a Markov chain on $\{0, 1, \ldots, 20\}$, where each state represents how long till the next bus. The transition probabilities are $p(0, k) = 1/20$ for each $1 \leq k \leq 20$ and $p(k, k-1) = 1$ for all $k > 0$.

*Item (ii):* The idea is that since the chain is stationary when you arrive your just sampling a state according to stationary distribution. Since each state represents the minutes until the next bus, the probability of waiting $i$ minutes for a bus is then just the stationary probability $\pi(i)$ of state $i$.

*Item (iii):* For the stationary distribution observe that

$$\begin{aligned}
\pi(0) &= \pi(1) \\
\pi(1) &= \frac{\pi(0)}{20} + \pi(2) \\
\pi(2) &= \frac{\pi(0)}{20} + \pi(3) \\
&\vdots \\
\pi(19) &= \frac{\pi(0)}{20} + \pi(20) \\
\pi(20) &= \frac{\pi(0)}{20}.
\end{aligned}$$

Thus we have $\pi(1) = \pi(1)/20 + \pi(2)$ thus since $\pi(0) = \pi(1)$,

$$\pi(2) = \frac{19\pi(0)}{20}.$$

However also we have $\pi(2) = \pi(0)/20 + \pi(3)$ thus $\pi(3) = 18\pi(0)/19$ and in general for any $1 \leq k \leq 20$,

$$\pi(k) = \frac{21-k}{20}\pi(0).$$

To find the values of $\pi$ all we must do is find $\pi(0)$, we do this by the fact $\pi$ is a probability distribution,

$$1 = \sum_{k=0}^{20} \pi(k) = \pi(0)\left(1 + \sum_{k=1}^{20} \frac{21-k}{20}\right),$$

which implies that $\pi(0) = 2/23$. Hence we have

$$\mathbb{E}_\pi[X_n] = \sum_{k=0}^{20} k\pi(k) = \sum_{k=1}^{20} k \frac{2}{23}\frac{21-k}{20} = \frac{154}{23} \approx 6.7 \text{ mins}.$$

*Item (iv):* The second moment is given by

$$\mathbb{E}_\pi[X_n^2] = \sum_{k=0}^{20} k^2\pi(k) = \sum_{k=1}^{20} k^2 \frac{2}{23}\frac{21-k}{20} = \frac{1617}{23}.$$

Thus the s.d. is

$$\sqrt{\frac{1617}{23} - \left(\frac{154}{23}\right)^2} = \frac{35\sqrt{11}}{23} \approx 5 \text{ mins}$$

---

**Question 3.** *Prove the following Lemma from class: For any probability distributions $\mu$ and $\eta$ on a countable state space $\Omega$*

$$\|\mu - \eta\|_{tv} = \frac{1}{2}\sum_{\omega\in\Omega} |\mu(\omega) - \eta(\omega)|.$$

---

*Solution:* Let $\Omega^+ = \{\omega : \mu(\omega) \geq \eta(\omega)\}$ and $\Omega^- = \{\omega : \mu(\omega) < \eta(\omega)\}$ . Then

$$\max_{A\subseteq\Omega} \mu(A) - \eta(A) = \mu(\Omega^+) - \eta(\Omega^+)$$

and

$$\max_{A\subseteq\Omega} \eta(A) - \mu(A) = \eta(\Omega^-) - \mu(\Omega^-).$$

Since $\Omega = \Omega^+ \cup \Omega^-$ and $\Omega^+ \cap \Omega^- = \emptyset$ we have

$$\mu(\Omega^+) + \mu(\Omega^-) = 1 \qquad \text{and} \qquad \eta(\Omega^+) + \eta(\Omega^-) = 1,$$

thus

$$\mu(\Omega^+) - \eta(\Omega^+) = \eta(\Omega^-) - \mu(\Omega^-).$$

Hence

$$\sup_{a\subset\Omega} |\mu(A) - \eta(A)| = |\mu(\Omega^+) - \eta(\Omega^+)| = |\mu(\Omega^-) - \eta(\Omega^-)|.$$

Combining the above yields

$$2\|\mu - \eta\|_{tv} = |\mu(\Omega^+) - \eta(\Omega^+)| + |\mu(\Omega^-) - \eta(\Omega^-)| = \sum_{\omega\in\Omega} |\mu(\omega) - \eta(\omega)|.$$

---

**Question 4.** *This question asks you to prove lower bounds on the mixing time of some lazy random walks on graphs.*

1. *Let $G = (V_1 \cup V_2, E)$ be a graph made of two disjoint complete graphs of $n$ vertices, supported respectively on $V_1$ and $V_2$, connected by a single edge. This is called the Barbell graph. Consider a lazy random walk on $G$. Prove that $t_{mix}(G) = \Omega(n^2)$ (recall from Lecture 8 that $t_{mix} = \tau(1/4)$).*

2. *Suppose now we add $s < n$ edges to the Barbell graph, where each edge has one endpoint in $V_1$ and the other endpoint in $V_2$. What happens to $t_{mix}(G)$?*

3. *Consider now a version of the Barbell graph where $|V_1| = n, |V_2| = \lfloor \log(n) \rfloor$ and there exists only an edge between $V_1$ and $V_2$. What is the mixing time of this graph?*

---

*Solution:* For part $(i)$: let $\pi$ be the stationary distribution of a lazy random walk in $G$ (recall that, for any vertex $u$, $\pi(u) = d(u)/2|E|$ where $d(u)$ is the degree of $u$). Now notice that, by symmetry, $\sum_{u\in V_1} \pi(u) = \sum_{u\in V_2} \pi(u) = 1/2$. You can prove this explicitly by using the formula for the stationary distribution mentioned above. Consider a probability distribution $p$ such that $\sum_{u\in V_2} p(u) \leq \epsilon$ for some small $\epsilon \geq 0$. Then,

$$\begin{aligned}
\|p-\pi\|_{TV} &= \frac{1}{2}\sum_{u\in V_1} |p(u)-\pi(u)| + \frac{1}{2}\sum_{u\in V_2} |p(u)-\pi(u)| \\
&\geq \frac{1}{2}\sum_{u\in V_1} (p(u)-\pi(u)) + \frac{1}{2}\sum_{u\in V_2} (\pi(u)-p(u)) \\
&= \frac{1}{2}\left(\sum_{u\in V_1} p(u) - \sum_{u\in V_1} \pi(u) + \sum_{u\in V_2} \pi(u) - \sum_{u\in V_2} p(u)\right) \\
&\geq \frac{1}{2}\left(1-\epsilon-\frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{2}-\epsilon\right) = \frac{1}{2}-\epsilon
\end{aligned}$$

where the last inequality follows from the facts that $\sum_{u\in V_2} p(u) \leq \epsilon$ and $\sum_{u\in V_1} \pi(u) = \sum_{u\in V_2} \pi(u) = 1/2$. Therefore, a walk to be mixed must have at least probability $\epsilon \geq 1/4$ to be in $V_2$.

But now suppose a walk start from a vertex $u \in V_1$ which is not the only vertex $v \in V_1$ adjacent to a vertex in $V_2$. Then, at each step, if the walk it's still in $V_1$, it has probability $O(1/n^2)$ to move to $V_2$ (because it must move first to $v$ and then move in $V_2$). Therefore, after $t$ steps, $\sum_{w\in V_2} P^t(u,w) = O(t/n^2)$ (this follows from a union bounds on the events "at step $i$ the walk moves from $V_1$ to $V_2$" for $i = 1,\ldots,t$). Hence, we need to wait $\Omega(n^2)$ before the walk is close to stationarity.

For part $(ii)$ repeat the same argument as in part $(i)$ but now at each step the probability to go from $V_1$ to $V_2$ is $\Omega(s/n^2)$. Therefore, $t_{mix} = O(n^2/s)$ (when you reach $V_2$, since the subgraph supported on $V_2$ is complete, after a few steps you are mixed).

For part $(ii)$, repeating again the same argument it is clear that to be mixed we just need to move from $V_2$ to $V_1$ (it is important here to notice that the worst case is to start in $V_2$: since $V_2$ is very small compared to $V_1$, if we start in the latter our argument doesn't work anymore). But this happens with probability $\Theta(1/(\log n)^2)$. Therefore mixing happens in $O(\log n)^2$ steps.

---

**Question 5.** *Let $\langle\cdot,\cdot\rangle_\pi$ be the inner product defined in the lecture. Show that it satisfies the following properties:*

Symmetry *For any $f, g \in \ell_2(\pi)$, $\langle f,g\rangle_\pi = \langle g,f\rangle_\pi$.*

Linearity *For any $f,g,h \in \ell_2(\pi)$ and $\alpha,\beta \in \mathbb{R}$, $\langle \alpha f+\beta g,h\rangle_\pi = \alpha\langle f,h\rangle_\pi + \beta\langle g,h\rangle_\pi$.*

Positive definiteness *For any $\underline{0} \neq f \in \ell_2(\pi)$, $\langle f,f\rangle_\pi > 0$.*

*Do all these properties hold if $\pi$ is not always positive?*

**Question 6.** *Given a matrix $M$ such that $Mf = \lambda f$ (i.e., $f$ is an eigenvector with eigenvalue $\lambda$ of $M$), prove that $M^k f = \lambda^k f$*

# Hints

**Hint** (Question 1)**.** *At face value $X_n$ is an (infinite) Markov chain on $\mathbb{N}$. We would like to consider it as a finite Markov chain, reduction $\bmod m$ (for some suitable $m$) will help us achieve this.*

**Hint** (Question 4(i))**.** *First prove that, for a distribution $p$ such that $\sum_{u\in V_2} p(u) \leq \epsilon$ for some small $\epsilon \geq 0$, $\|p-\pi\|_{TV} \geq \frac{1}{2}-\epsilon$, where $\pi$ is the stationary distribution of a random walk on the Barbell graph. Use this fact to obtain a lower bound on the mixing time (think how many steps we need so that $\sum_{u\in V_2} p^t(u) > \epsilon$, where $p^t$ is the random walk distribution after $t$ steps).*