Probability and Computation: Problem sheet 1 Solutions

Question 1. *Recall that a permutation $\sigma : [n] \to [n]$ is a bijection from $[n]$ to $[n]$. A cycle $c = (c_1, c_2, \dots, c_k)$ is a sequence such that $\sigma(a_i) = a_{i+1 \mod k}$. Example: if $\sigma(1) = 3$, $\sigma(2) = 2$, $\sigma(3) = 4$ and $\sigma(4) = 1$ (this is also denoted $\sigma = \binom{1\,2\,3\,4}{3\,2\,4\,1}$) then $(1, 3, 4)$ is a cycle (as are $(3, 4, 1)$ and $(4, 1, 3)$) .*

(i) How many permutations on $[n]$ are there?

(ii) Show that the number of permutations on $[n]$ which contain a cycle of length $\ell > \lfloor n/2 \rfloor + 1$ is $n!/\ell$.

(iii) Let $\mathcal{E}$ be the event that a uniformly distributed random permutation contains no cycle longer than $\lfloor n/2 \rfloor$. Show that $\mathbf{P}[\,\mathcal{E}\,] = 1 - \left(\frac{1}{\lfloor n/2 \rfloor + 1} + \ldots + \frac{1}{n}\right)$.

(iv) Approximate the expression for $\mathbf{P}[\,\mathcal{E}\,]$ above by integrals to give $\mathbf{P}[\,\mathcal{E}\,] = 1 - \ln 2 + o(1) \approx 0.31183$.

Solution: Part (*i*) there are $n!$ permutations on $[n]$.

For part (*ii*) each permutation on $[n]$ can contain at most one cycle longer than $\lfloor n/2 \rfloor + 1$ and there are exactly $\binom{n}{\ell}$ ways to select the elements of such a cycle. Within this cycle, the elements can be arranged in $\ell!/\ell = (\ell - 1)!$ ways since two cycles are the same if one is a rotation of the other. The remaining elements of the permutation can be arranged in $(n - \ell)!$ ways. Therefore, the number of permutations on $[n]$ with a cycle of length ℓ is equal to

$$\binom{n}{\ell} \cdot (\ell - 1)! \cdot (n - \ell)! = \frac{n!}{\ell}.$$

For part (*iii*) a permutation of the numbers 1 to n can contain at most one cycle of length $\ell > \lfloor n/2 \rfloor$. Since the event that a uniform permutation contains a cycle of length $\ell > \lfloor n/2 \rfloor$ are disjoint their probability follows by summing the result of part (*ii*) and dividing by the total number of permutations ($n!$). Thus we have

$$\begin{aligned}\mathbf{P}[\,\text{No cycle } \geq \lfloor n/2 \rfloor\,] &= 1 - \sum_{\ell = \lfloor n/2 \rfloor}^{n} \mathbf{P}[\,\text{longest cycle } = \ell\,] = 1 - \frac{1}{n!}\left(\frac{n!}{\lfloor n/2 \rfloor + 1} + \cdots + \frac{n!}{n}\right) \\ &= 1 - \left(\frac{1}{\lfloor n/2 \rfloor + 1} + \ldots + \frac{1}{n}\right).\end{aligned}$$

For part (*iv*) the expression above is $1 - H_n - H_{\lfloor n/2 \rfloor + 1}$ where H_n is the n-th harmonic number. Observe that for any $n \geq 2$

$$\int_{\lfloor n/2 \rfloor + 1}^{n+1} \frac{1}{x} dx \leq \frac{1}{\lfloor n/2 \rfloor + 1} + \ldots + \frac{1}{n} \leq \int_{\lfloor n/2 \rfloor}^{n} \frac{1}{x} dx,$$

where the two integrals above evaluate to $\ln 2 \pm \ln(1 - O(1/n))$. The result follows by Taylor's approximation for the natural logarithm $\ln(\cdot)$.

Question 2. *Recall that a probability vector (distribution) is a non-negative real vector whose elements sum to 1. A stochastic matrix is a real square matrix, where each row is a probability vector. Observe every Stochastic matrix gives rise to a Markov chain and visa versa.*

(i) Let $\nu \in \mathbb{R}_+^n$ be a probability vector and $M \in \mathbb{R}_+^{n \times n}$ be a stochastic matrix. Show that νM is a probability vector.

A doubly stochastic matrix is a real square matrix, where each row and column is a probability vector.

(ii) Prove that the uniform distribution is stationary for any Markov chain whose transition matrix is doubly stochastic.

Question 3. *In this question we shall consider two probability measures μ_1, μ_2 on the common state space $[12] = \{1, \dots, 12\}$. The first measure is the uniform measure $\mu_1(x) = 1/12$ for all $x \in [12]$, this is a "12 sided fair die". The second μ_2 is the measure generated by the sum of two independent, fair, 6-sided die i.e. $\mu_2(1) = 0, \mu_2(2) = 1/6^2, \dots$. Calculate $\|\mu_1 - \mu_2\|$.*

Question 4. *Recall Jerry's Coupling Z_t for two Simple Random walks T_t, J_t on the 3-cycle. This coupling is given by Markov chain $Z_t = (T_t, J_t)$ on $\{0,1,2\} \times \{0,1,2\}$:*

- *Run the Cat T_t as normal*
- *Mouse J_t moves according to the rule:* $J_{t+1} = \begin{cases} J_t + 1 \mod 3 & \textit{if } T_{t+1} = T_t + 1 \mod 3 \\ J_t - 1 \mod 3 & \textit{if } T_{t+1} = T_t - 1 \mod 3. \end{cases}$

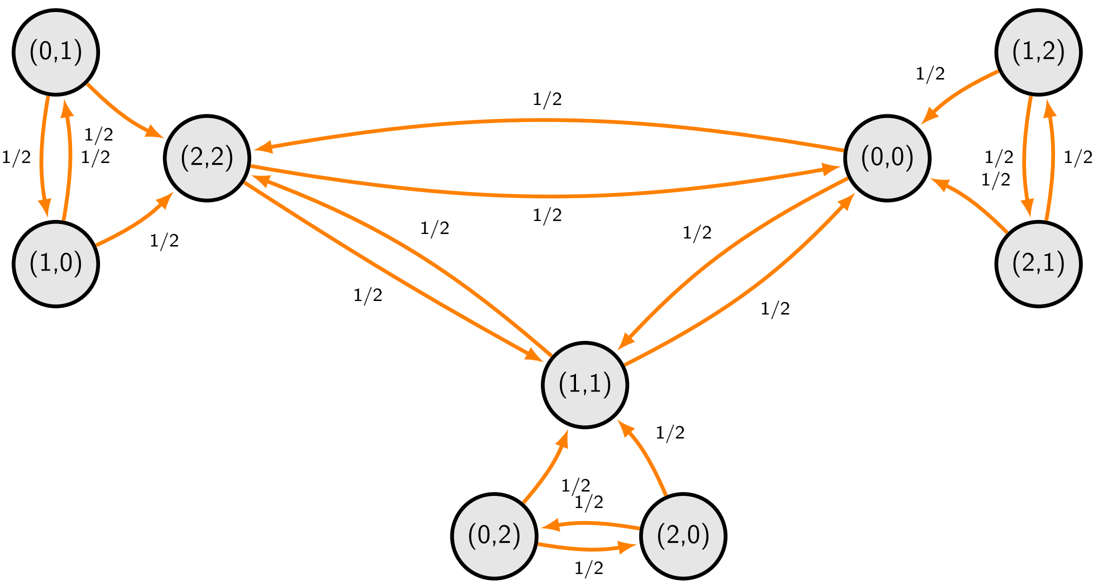
This coupling lets the "Mouse" J_t avoid the "Cat" T_t forever.

1. *Write down the rule for "Tom's Coupling" - the coupling where T_t and J_t meet as soon as possible and show this is a valid coupling.*
2. *Draw a diagram of the Markov Chain Z_t on $\{0,1,2\} \times \{0,1,2\}$ generated by Tom's Coupling.*

Solution: Clearly the Mouse's marginal distribution is correct, what about Cat? Let $\mathbf{P}_{(x,y)}[\cdot] := \mathbf{P}[\cdot \mid Z_t = (x,y)]$, then for $x, y, z \in \{0,1,2\}$

$$\begin{aligned} \mathbf{P}_{(x,y)}[T_{t+1} = x] &= \mathbf{P}[T_{t+1} = x, J_t = x] + \mathbf{P}[T_{t+1} = x, J_t = z] = 0 + 0 = 0 \\ \mathbf{P}_{(x,y)}[T_{t+1} = y] &= \mathbf{P}[T_{t+1} = y, J_t = x] + \mathbf{P}[T_{t+1} = y, J_t = z] = 1/2 + 0 = 1/2 \\ \mathbf{P}_{(x,y)}[T_{t+1} = z] &= \mathbf{P}[T_{t+1} = z, J_t = x] + \mathbf{P}[T_{t+1} = z, J_t = z] = 0 + 1/2 = 1/2 \end{aligned}$$

Thus the Tom has the right transition densities. Z_t looks like



Question 5. *In lecture 1 we saw the balls in bins experiment where one assigns m balls to n bins uniformly and independently. By considering the number of balls assigned to bins between the first time i bins are empty and $i-1$ bins are empty, show that the the expected number of balls one must assign before there is no empty bin is $n \log n + \Theta(n)$.*

Solution: Let T_i be number of balls assigned while you have exactly i bins still empty. The r.v. T_i is geometrically distributed $Geo(p_i)$ for some p_i.

There are i bins empty out of n total and all bins are equally likely thus

$$p_i = i/n.$$

The total number T of balls assigned is given by $\sum_{i=1}^{n} T_i$ and so

$$\mathbf{E}[\,T\,] = \sum_{i=1}^{n} \mathbf{E}[\,T_i\,] = \sum_{i=1}^{n} \frac{n}{i} = n \cdot H_n = n \log n + \Theta(n)\,,$$

where H_n is the n-th harmonic number.

Question 6. *Let X_n be the sum of n independent rolls of a fair die. Show that, for any $k \geq 2$,*

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

Hint. *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 $\mod m$ (for some suitable m) will help us achieve this.*

Question 7. *State j is accessible from state i if, for some integer $n \geq 0$, $P^n_{i,j} > 0$. If two states i and j are accessible from each other, we say that they communicate and we write $i \sim j$. Prove that communicating relation $\sim$ defines an equivalence relation.*

Question 8. *Prove the following Lemma from class: For any probability distributions μ and η on a countable state space Ω*

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

Hint. *Recall the sets $\Omega^{\pm}$ from the Coupling Lemma.*

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 9. *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 3 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 π 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.
