Probability and Computation: Problem sheet 2 Solutions

Question 1. *A Lion and a Deer each independently take a random walk on a connected, undirected, non-bipartite graph G. They start at the same time on different nodes, and each makes one transition at each time step. The Lion eats the Deer if they are ever at the same node at some time step. Let n and m denote, respectively, the number of vertices and edges of G. Show an upper bound of $O(m^2n)$ on the expected time before the cat eats the mouse.*

Solution: If we consider the motion of both animals simultaneously then we have a Markov chain Z_t on $V \times V$, where a state is a pair (l, d) where the first position indicates the Lion's position and the second indicates the Deer's position, and transition probabilities given by

$$\mathbf{P}[\, Z_{t+1} = (c, d) \mid Z_t = (a, b)\,] = \frac{1}{d(a)\, d(b)}.$$

Notice that this is a simple random walk on a new graph $\tilde{G}$ with $\tilde{V} = V \times V$ where two states $(a, b), (c, d)$ are connected by an edge in $\tilde{E}$ if $ac \in E$ and $bd \in E$.

Observe that if $Z_t = (x, x)$ for any $x \in V$ then the two animals will meet. Thus we fix any state (x, x) and bound the hitting time $\tilde{h}_{(l,d),(x,x)}$ from above by $O(m^2n)$ for any starting state (l, d).

Since the graph is non-bipartite there are paths of odd and even lengths, each of length at most $3n$ (certainly less but this is fine), from l to x and from d to x. Also there is a path of length 2 from x back to x. Choose two of the paths, one from l to x and the other from d to x, both of the same parity. W.l.o.g. the path from l to x is shortest and so have the lion do loops of length 2 from x to x while waiting for the Deer. This way we have constructed a path P of length at most $3n$ from (l, d) to (x, x) in $\tilde{G}$. By the bound on hitting times given in the slides, if $(x, y)(w, z) \in \tilde{E}$ then $\tilde{h}_{(x,y)(w,z)} \le 2|\tilde{E}|$. However by the handshake lemma

$$2|\tilde{E}| = \sum_{v \in \tilde{E}} d(v) = \sum_{(x,y) \in V \times V} d(x)\, d(y) = 4m^2.$$

Thus it follow by summing hitting times along the path P from (l, d) to (x, x) that

$$\tilde{h}_{(l,d)(x,x)} \le \sum_{(x,y)(w,z) \in P} \tilde{h}_{(x,y)(w,z)} \le 3n \cdot 2|\tilde{E}| = O(m^2n)\,.$$

Question 2. *Let $t_{hit}(G) = \max_{x,y \in V} \mathbf{E}_x\left[\tau_y^+\right]$ be the maximum hitting time. Prove the following weak version of the Matthew's bound: For any graph G we have $t_{cov}\,(\,G\,) = O(t_{hit} \cdot \log n)\,.$*

Solution: Divide $6t_{hit}\,(\,G\,)\log_2 n$ random walk steps into $3\log_2 n$ epochs each of length $2t_{hit}\,(\,G\,)$. For any x, y

$$\mathbf{P}_x\left[\tau_y^+ \ge 2t_{hit}\right] \overset{\text{Markov Ineq.}}{\le} \mathbf{E}_x\left[\tau_y^+\right]/2t_{hit} \le 1/2.$$

Hence for any x, y,

$$\mathbf{P}_x\left[\tau_y^+ \ge 6t_{hit} \cdot \log_2 n\right] \overset{\text{Markov Prop.}}{\le} \left(\max_{z \in V} \mathbf{P}_z\left[\tau_y^+ \ge 2t_{hit}\right]\right)^{3\log_2 n} \le 2^{-3\log_2 n} = n^{-3}.$$

Let $\mathcal{E}$ be the event $\{\tau_{cov} < 6t_{hit}\log_2 n\}$. Thus we have

$$\mathbf{P}[\,\mathcal{E}^c\,] = \max_{z \in V} \mathbf{P}_z\left[\cup_{y \in V}\{\tau_y^+ \ge 6t_{hit} \cdot \log_2 n\}\right] \overset{\text{Union Bound}}{\le} n \cdot n^{-3} = n^{-2}.$$

If $\mathcal{E}^c$ occurs we use our previous upper bound of $4|V| \cdot |E| \leq 2n^3$. Thus

$$\begin{aligned} t_{cov}(G) &\leq \max_{z\in V} \mathbf{E}_z[\tau_{cov} \mid \mathcal{E}]\, \mathbf{P}[\mathcal{E}] + \max_{z\in V} \mathbf{E}_z[\tau_{cov} \mid \mathcal{E}^c]\, \mathbf{P}[\mathcal{E}^c] \\ &\leq 6t_{hit}(G)\log n + 2n^3 \cdot n^{-2} = O(t_{hit} \cdot \log n). \end{aligned}$$

□

Question 3. *Let X and Y be two Binomial random variables with parameters (n, p_1) and (n, p_2) respectively, for a positive integer n and for $0 \leq p_1 \leq p_2 \leq 1$. Show that for any k*

$$\mathbf{P}[X \geq k] \leq \mathbf{P}[Y \geq k].$$

(i) Define a function $f(p) = \mathbf{P}[B(n,p) \geq k]$ and show that it is increasing in p.

(ii) Now use coupling to prove the same statement, extending the coupling of two unfair coins given in the lecture.

Solution: First, we can show it using a simple computation. We need to show that the function

$$f(p) = \mathbf{P}[B(n,p) \geq k] = \sum_{i=k}^{n} \binom{n}{i} p^i (1-p)^{n-i}$$

is increasing i.e. it is sufficient to show that $f'(p) \geq 0$.

$$\begin{aligned} f'(p) &= \sum_{i=k}^{n} \binom{n}{i} \left(ip^{i-1}(1-p)^{n-i} - (n-i)p^i(1-p)^{n-i-1}\right) \\ &= \sum_{i=k}^{n} \binom{n}{i} (i(1-p) - (n-i)p)p^{i-1}(1-p)^{n-i-1} \\ &= \sum_{i=k}^{n} \binom{n}{i} (i-np)p^{i-1}(1-p)^{n-i-1} \\ &\geq \sum_{i=0}^{n} \binom{n}{i} (i-np)p^{i-1}(1-p)^{n-i-1} \\ &= \sum_{i=0}^{n} \binom{n}{i} ip^{i-1}(1-p)^{n-i-1} - np\sum_{i=0}^{n} \binom{n}{i} ip^{i-1}(1-p)^{n-i-1} \\ &= \frac{n}{1-p}\sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} - \frac{n}{1-p}\sum_{i=0}^{n} \binom{n}{i} p^{i}(1-p)^{n-i} \\ &= \frac{n}{1-p}\left((p+(1-p))^{n-1} - ((p+(1-p))^n\right) \\ &= 0, \end{aligned}$$

where for the inequality above we used the fact that the terms in the sum are negative only for $i < np$, so we can lower bound it by taking all summands.

Now we present a more elegant solution that uses coupling. X and Y can be both represented as a sum of n Bernoulli random variables with parameters p_1 and p_2 respectively.

$$X = \sum_{i=1}^{n} X_i \text{ and } Y = \sum_{i=1}^{n} Y_i$$

We couple X_i and Y_i as follows. Let $Z_i = \mathit{Bernoulli}\left(\frac{p_2-p_1}{1-p_1}\right)$. Then let $X_i = \mathit{Bernoulli}(p_1)$ and $Y_i = \max\{X_i, Z_i\} \geq X_i$. It is not hard to see that this is a valid coupling and it produces a valid coupling of X and Y. Since $X \leq Y$ the statement of the problem holds trivially.

Question 4. *Consider an infinite ladder with steps numbered starting from 0. When at step n, we move to step $(n+1)$ with probability p_n or fall to step 0 with probability $1-p_n$. The sequence p_n is increasing (we are being more and more cautious). Prove that $\mathbf{P}[\, X_n > k | X_0 = j\,]$ is an increasing function of j. Also, show that the condition that p_n is increasing is necessary.*

Solution: Solution given in problem class, similar to the above.

Question 5. *We saw the Top-to-Random shuffle in lectures. One can also shuffle cards by doing the reverse of this process, the Random-to-Top Shuffle: at each time step a card is taken at random from the deck and placed on the top.*

(i) Construct a coupling between two copies of the R-to-T chain so that after some (random) time they will be in the same state.

(ii) Apply the Coupling Lemma for Mixing to your coupling to bound the mixing time of the R-to-T Chain.

Solution: See Mitzenmacher and Upfal, Section 12.2.1.

Question 6. *Consider the Balls-into-Bins setting where we have n balls and n bins. We assume that n is large enough. We are going to prove that the maximum load is whp. at least $c \log n / \log\log n$ for some $c > 0$.*

(i) Let $Y_j(k)$ be the random variable that indicates that bin j receives at least k balls. Prove that for any $k \le n$, it holds that

$$\mathbf{P}[\, Y_j(k) = 1\,] \ge \frac{e^{-2}}{k^k}.$$

(ii) Show it exist $c > 0$ such that for $k^* = \left\lfloor \frac{c \log n}{\log\log n} \right\rfloor$, we have

$$\mathbf{P}[\, Y_j(k^*) = 1\,] \ge n^{-1/3},$$

(iii) Argue that for any $k \le n$, and any bins i, j we have

$$\mathbf{P}[\, Y_j(k) Y_i(k) = 1\,] \le \mathbf{P}[\, Y_i(k) = 1\,]\, \mathbf{P}[\, Y_j(k) = 1\,]$$

(iv) Let $Y = \sum_{j=1}^{n} Y_j(k^*)$. Check that $\mathbf{E}[\, Y\,] \ge n^{2/3}$ and that $\mathbf{Var}\,[\, Y\,] \le n$

(v) Conclude that $\mathbf{P}[\, Y = 0\,] \le n^{-1/3}$

Solution: Solution given in problem class.

Hint (Collected hints for the exercises)**.**

Q1: Consider a Markov chain whose states are the ordered pairs (a, b), where a is the position of the Lion and b is the position of the Deer.

Q2: Use Markov's Inequality.

Q3: Hint 1: $\sum_{i=k}^{n} a_i \ge \sum_{i=0}^{n} a_i$ if the negative terms of the sum are all in the beginning.
Hint 2: $\sum_{i=0}^{n} \binom{n}{i} i p^{i-1} (1-p)^{n-i-1} = 0$. Why?

Q5(i): Run one copy of the chain X_t as normal, for the other make sure that at each time step the same card as in X_t is on the top.

Q5(iii): Similar to Balls and bins makes an appearance here.

Q6(i): Hint 1: Compute of getting exactly k balls in the bin.
Hint 2: Use that $1 - x \ge e^{-2x}$ for $x \in [0, 1/2]$ and that $\binom{n}{k} \ge (n/k)^k$.