Probability and Computation: Mock Exam Solutions

Question 1. *Consider the balls into bins problem where m balls are assigned uniformly and independently at random to n bins, where $m > n$. Let X be the number of empty bins.*

(a) Compute $\mathbf{E}[\,X\,]$.

(b) Prove that X is Liptschitz as a function of the bin number to which each ball is assigned.

(c) Use McDiarmid's inequality to derive an upper bound for $\mathbf{P}[\,X > \mathbf{E}[\,X\,] + t\,]$ provided $t > 0$.

(d) Find a better bound for $\mathbf{P}[\,X > \mathbf{E}[\,X\,] + t\,]$ by expressing X as a function of something different.

Solution:

(a) A bin is empty if no balls are assigned to it. This event occurs with probability $(1-1/n)^m$. Therefore, the expected number of empty bins is $n(1-1/n)^m$.

(b) Let $Y_1, \dots, Y_m$ be the location of the ball, so ball $i \in [1, m]$ is placed in bin number $Y_m \in [1, n]$. The number of empty bins X can be expressed as a function $X = f(Y_1, \dots, Y_m)$ of the ball locations. Note that changing the value of Y_i increases or reduces the number of empty bins in at most 1. Thus X is Liptschitz with constant $c = (1, 1, \dots, 1)$.

(c) Since the location of each ball is independent, McDiarmid's inequality yields

$$\mathbf{P}[\,X > \mathbf{E}[\,X\,] + t\,] \le \exp\left(-\frac{2t^2}{c_1^2 + \cdots + c_m^2}\right) = \exp\left(-2t^2/m\right) \tag{1}$$

(d) This question is hard: one will be tempted to write this as a function of the number of empty bins, but that does not work as they are not independent. So we need to go a bit deeper.

Define the Doob martingale $X_t = \mathbf{E}[\,X|Y_1, \dots, Y_t\,]$ and $X_0 = \mathbf{E}[\,X\,]$. We may think of X_t as the process where we assign the m balls one by one, and after assigning a ball ask: what is the expected number of empty bins at the end of the process? Recall that $X_m = \mathbf{E}[\,X|Y_1, \dots, Y_m\,] = X$.

We are going to use the Azuma-Hoeffding inequality (Lecture 7). To apply this, we need to find a_t and b_t such that

$$a_t \le X_t - X_{t-1} \le b_t.$$

Let c_t be the number empty bins after we assign the first t balls (which clearly is a function of $Y_1, \dots, Y_t$). Then we can explicitly compute X_t. Consider a bin that is part of the c_t empty bins. Then the probability that this bin is empty after an additional $m-t$ balls is $(1-1/n)^{m-t}$ Therefore

$$X_m = \mathbf{E}\left[\sum_{i=1}^{c_t} \mathbf{1}_{Y_{t+1}\neq i, \dots, Y_m \neq i} | Y_1, \dots, Y_t\right] = \sum_{i=1}^{c_t} \mathbf{E}\left[\,\mathbf{1}_{Y_{t+1}\neq i, \dots, Y_m \neq i} | Y_1, \dots, Y_t\,\right] = c_t(1-1/n)^{m-t},$$

by independence of ball locations. Our next step is to compute (bound) the differences

$$X_t - X_{t-1} = c_t(1-1/n)^{m-t} - c_{t-1}(1-1/n)^{m-t+1}.$$

Note that c_t takes values c_{t-1} or $c_{t-1} - 1$ (either the t-th ball is assigned to an empty bin or not), thus

$$X_t - X_{t-1} \le c_{t-1}(1-1/n)^{m-t} - c_{t-1}(1-1/n)^{m-t+1} = \frac{c_{t-1}}{n}(1-1/n)^{m-t}$$

and

$$X_t - X_{t-1} \geq (c_{t-1} - 1)(1-1/n)^{m-t} - c_{t-1}(1-1/n)^{m-t+1} = \frac{c_{t-1}}{n}(1-1/n)^{m-t} - (1-1/n)^{m-t}.$$

Hence $b_t - a_t \leq \left(1 - \frac{1}{n}\right)^{m-t}$ and $\sum_{i=1}^m (b_i - a_i)^2$ equals

$$\sum_{i=1}^{m} \left(\left(1 - \frac{1}{n}\right)^2\right)^{m-i} = \sum_{i=0}^{m-1} \left(\left(1 - \frac{1}{n}\right)^2\right)^{i} \leq \sum_{i=0}^{\infty} \left(\left(1 - \frac{1}{n}\right)^2\right)^{i} = \frac{1}{1 - \left(1 - \frac{1}{n}\right)^2} = \frac{n}{2 - 1/n}.$$

The Azuma-Hoeffding inequality tells us that

$$\mathbf{P}[\, X_m - X_0 \geq t \,] \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^m (b_i - a_i)^2}\right) \leq \exp\left(-\frac{2t^2\,(2 - 1/n)}{n}\right)$$

and similarly for $\mathbf{P}[\, X_m - X_0 \leq -t \,]$. This bound is better than $\exp\left(-\frac{2t^2}{m}\right)$ obtained in part (b).

Question 2.

(a) *Fill the missing entries in the matrix below so that it represents the transition matrix of a reversible Markov chain:*

$$\begin{pmatrix} 0 & 3/4 & 1/4 & 0 & \dots \\ \dots & 0 & \dots & 0 & 0 \\ \dots & 4/7 & 0 & 2/7 & 0 \\ 0 & 0 & \dots & 0 & \dots \\ 0 & 0 & \dots & \dots & 0 \end{pmatrix}$$

(b) *Find the stationary distribution of your matrix. Is the corresponding Markov chain irreducible? Is it aperiodic? Explain your answers.*

(c) *Let P be a transition matrix of a Markov chain on state space Ω. Let π be a probability distribution satisfying the following equation:*

$$\pi(x)P(x,y) = \pi(y)P(y,x) \quad \forall x, y \in \Omega.$$

Prove that π is a stationary distribution for P.

Solution:

(a) To come up with a solution, recall that, for P to be a transition matrix, each row must sum to one. Moreover, to enforce reversibility, you should make sure $P(x,y) = 0 \iff P(y,x) = 0$. To fill the entries that you cannot fill up enforcing these two conditions, remember that P is reversible if and only if can be represented as a transition matrix of an undirected weighted graph. This is a possible solution:

$$\begin{pmatrix} 0 & 3/4 & 1/4 & 0 & 0 \\ 3/7 & 0 & 4/7 & 0 & 0 \\ 1/7 & 4/7 & 0 & 2/7 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

It is a transition matrix because each row sum to 1. It is reversible because it represents the transition matrix of a random walk on the weighted graph $G = (V, E, w)$ defined as follows: $V = \{1, \dots, 5\}, E = \{\{1,2\}, \{1,3\}, \{2,3\}, \{3,4\}, \{4,5\}\}$, with corresponding weights $\{3, 1, 4, 2, 2\}$.

(b) The stationary distribution of a random walk on an undirected graph has form $\pi(u) = d(u)/\sum_z d(z)$. In this case $(1/6, 7/24, 7/24, 1/6, 1/12)$. The Markov chain is irreducible since the graph is connected, and it is aperiodic since the graph is not bipartite (remember this only works for random walks on undirected graphs).

(c) We want to prove that $\pi P = \pi$. For any $x \in \Omega$ we have that

$$(\pi P)(x) = \sum_{y \in \Omega} \pi(y) P(y,x) = \sum_{y \in \Omega} \pi(x) P(x,y) = \pi(x) \sum_{y \in \Omega} P(x,y) = \pi(x),$$

where the last line follows because each row of P sums to 1.

Question 3. *A matching in a graph is a set of edges without common vertices. In the Maximum Bipartite Matching problem, we are given a bipartite graph $G(L \cup R, E)$ (without multiple edges), and we want to find a matching of maximum cardinality. Consider the following randomised algorithm for this problem: Each edge is selected independently with probability p. All edges that have common endpoints are discarded. Assume that the bipartite graph has $|L| = |R| = n$ and that every vertex has degree 3.*

(a) What is the expected cardinality of the matching returned by the algorithm as a function of p?

(b) Find the value of p that maximises the expected cardinality of the matching. What is the expected cardinality of the matching in this case?

(c) Assume now the graph is regular of degree $d \geq 3$, not necessarily constant. Would you choose a constant value of p or a value that depends on d and/or n? Explain your choice.

Solution:

(a) Let $M \subseteq E$ be the random set of edges included in the matching. Then by linearity of expectations,

$$\mathbf{E}[\,|M|\,] = \mathbf{E}\left[\sum_{\{u,v\} \in E} \mathbf{1}_{\{u,v\} \in M} \right] = \sum_{\{u,v\} \in E} \mathbf{E}\big[\, \mathbf{1}_{\{u,v\} \in M} \,\big] = \sum_{\{u,v\} \in E} \mathbf{P}[\, \{u,v\} \in M \,]$$

Let us now consider the event $\{u,v\} \in M$ for a fixed edge $\{u,v\} \in E(G)$. With probability p, the edge is selected in the first phase. Conditional on this, with probability $(1-p)^2 \cdot (1-p)^2 = (1-p)^4$ the other two incident edges to u and v are not included. Thus with probability

$$p \cdot (1-p)^4$$

any fixed edge $\{u,v\} \in E(G)$ is included. Thus

$$\mathbf{E}[\,|M|\,] = p \cdot (1-p)^4 \cdot |E|. \tag{2}$$

(b) Let us denote

$$f(p) := p \cdot (1-p)^4.$$

Hence,

$$f'(p) = (1-p)^4 + -4p \cdot (1-p)^3 = (1-p)^3 \cdot (1-5p), \tag{3}$$

and

$$f''(p) = 4(5p-2) \cdot (1-p)^2 \tag{4}$$

There are two roots $p = 1$ and $p = 1/5$ to (3), i.e. p such that $f'(p) = 0$, we must check both to see if either is a maximum. Firstly plugging $p = 1$ into (4) yields $f''(1) = 0$, similarly for $p = 1/5$ we have $f''(1/5) = 4(-1)(4/5)^2 < 0$. Thus $p = 1/5$ is a maxima for $f(p)$, using this in the expression (2) for the expected size of the matching m yields

$$\mathbf{E}[\,|M|\,] = (1/5) \cdot (4/5)^4 \cdot |E| = 2^8 \cdot |E|/5^5.$$

(c) The following is a rigorous derivation, given for the sake of completeness, but not needed for full marks. It is also sufficient to say that, p needs to decrease as d grows to infinity. The formula from part a) extended to arbitrary d demonstrates that the optimal value of p should not depend on n.

Rigorous Derivation: In the case where G is d-regular,

$$\mathbf{E}[\,|M|\,] = p \cdot (1-p)^{2(d-1)} \cdot |E| =: f_d(p).$$

Thus

$$f'_d(p) = (1-p)^{2d-3} \cdot (1 - 2dp + p),$$

and

$$f''_d(p) = 2 \cdot (d-1) \cdot (1-p)^{2(d-2)} \cdot ((2 \cdot d - 1)p - 2)$$

Hence setting $1 - 2dp + p = 0$ yields $p(1-2d) = -1$ and thus $p = \frac{1}{2d-1}$. Also it can be seen that for this choice of p, $f''_d(p) < 0$, thus it is a local maximum. Thus the optimal choice of p is independent of n.