Probabilistic Techniques

Problem set #5 — Markov chains and Balls&Bins

Assignment:	18. 12. 2019
Hints:	For individual problems, upon request after 9. 1. 2020 (10 pushups for a hint :D)
Deadline:	31. 1. 2020

Note that some of the problems might require material from the last lecture (9. 1. 2019).

1. By a *homogeneous discrete time Markov chain* we mean a sequence of random variables $X_0, X_1, \dots$ such that for every n we have $\Pr[X_{n+1} = k | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0] = \Pr[X_{n+1} = k | X_n = i]$.

   (a) Let $Y_0, Y_1, \dots$ be integers chosen independently uniformly from $\{-1, 0, 1\}$ and let $X_n = \max\{Y_0, \dots, Y_n\}$. Decide whether $X_0, X_1, \dots$ is a homogeneous discrete time Markov chain. **[1 point]**

   (b) Let $Y_0, Y_1, \dots$ be integers chosen independently uniformly from $\{-1, 0, 1\}$ and let $X_n = \max\{Y_n, Y_{n+1}\}$. Decide whether $X_0, X_1, \dots$ is a homogeneous discrete time Markov chain. **[1 point]**

2. At the exam, Chad was given the following task: "Count to k." Chad starts with $X = 0$ and for n steps does the following:

   - With probability 1/4 decides that $X < k$ and therefore increments X by 1,
   - with probability 1/4 decides that $X > k$ and therefore decrements X by 1, and
   - with probability 1/2 decides that he needs more time to think and does nothing (which still counts as a step).

   Prove that there are constants $0 < c_1, c_2$ (independent from n) such that $c_1\sqrt{n} \leq \mathbb{E}[|X|] \leq c_2\sqrt{n}$. **[2 points]**

3. Construct satisfiable 2-SAT formulas with n variables (for arbitrarily large n) such that the randomized 2-SAT algorithm from the lecture takes $\Omega(n^2)$ steps (in expectation) to find a satisfying assignment. (Recall that in each step, the algorithm picks an arbitrary unsatisfied clause and flips one of its variables.) **[2 points]**

4. We say that a graph G is triangle 2-colourable if there is a 2-colouring of the vertices of G such that no triangle in G is monochromatic. (Convince yourself that every 3-colourable graph is triangle 2-colourable.)

   Let G be a 3-colourable graph. Start with $c\colon V(G) \to \{0, 1\}$ being constant 0 and then do the following. While there is a monochromatic triangle in G (with respect to c), pick one of its vertices at random and flip its colour. Show that the expected number of steps for this process to find a suitable triangle 2-colouring is polynomial in $|V(G)|$. **[3 points]**

5. Let X be a Poisson random variable with mean $\mu \in \mathbb{Z}$. Prove that $\Pr[X \geq \mu] \geq \frac{1}{2}$. (Hint: Show that $\Pr[X = \mu + h] \geq \Pr[X = \mu - h - 1]$ for every $0 \leq h \leq \mu - 1$.) **[2 points]**

6. Throw m balls uniformly independently into n bins, where the bins are numbered from 0 to $n - 1$. We say that there is a *k-gap* starting at j if bins j, $j + 1$, …, $j + k - 1$ are all empty.

   (a) Compute the expected number of k-gaps. **[1 point]**

   (b) Prove a Chernoff-like bound for the number of k-gaps X, that is, the inequality

   $$\Pr[X \geq (1 + \delta)\mathbb{E}[X]] \leq e^{-c\mathbb{E}[X]\delta^2},$$

   where $0 < \delta < 1$ and c is a positive constant. **[4 points]**

7. Consider the following variant of balls into bins, where the bins are numbered from 0 to $n - 1$. Let us have $k = \log_2 n$ batches of balls, each containing $b = \lceil n/\log_2 n \rceil$ balls. The process will consist of k rounds. In each round, we choose starting bin j uniformly at random and put one ball into each of the bins numbered $j, (j+1) \bmod n, \dots, (j + b - 1) \bmod n$. Show that the maximum load in this case is only $O(\log\log n/\log\log\log n)$ with probability approaching 1 as n goes to infinity. **[3 points]**