Partner 1
Partner 2
Partner 3

Homework 3

CSCI 1550 / 2540
April 18th, 2024

Due: April 18th, 2024
Remember to show your work for each problem to receive full credit.

Problem 1 (20 points)

A parking-lot attendant has mixed up n keys for n cars. The n car owners arrive together. The attendant gives each owner a key according to a permutation chosen uniformly at random from all permutations. If an owner receives the key to their own car, they take it and leave; otherwise, they return the key to the attendant. The attendant now repeats the process with the remaining keys and car owners. This continues until all owners receive the keys to their cars.

Let R be the number of rounds until all car owners receive the keys to their cars. We want to compute $\mathbb{E}[R]$. Let X_i be the number of owners who receive their car keys in the ith round.

1. Prove that
$$Y_i = \sum_{j=1}^{i} (X_j - \mathbb{E}(X_j \mid X_1, \ldots, X_{j-1}))$$
is a martingale with respect to the X_i's.

2. Use the martingale stopping theorem to compute $\mathbb{E}[R]$.

Problem 2 (20 points)

Consider a random walk on the infinite two dimension integer grid:

$$G = \{(x, y) \mid x \in \{-\infty, \infty\},\ x \in \{-\infty, \infty\}\}.$$

The random walk starts at $(0, 0)$, and if the walk is at (x_t, y_t) at time t, then with equal probabilities the walk moves to one of the adjacent nodes $(x_t - 1, y_t)$, $(x_t, y_t - 1)$, $(x_t + 1, y)$, or $(x_t, y_t + 1)$. I.e.

$$Pr((x_{t+1}, y_{t+1})) \mid (x_t, y_t)) = \begin{cases} 1/4 & \text{if} \quad (x_{t+1}, y_{t+1}) = (x_t - 1, y_t) \\ 1/4 & \text{if} \quad (x_{t+1}, y_{t+1}) = (x_t + 1, y_t) \\ 1/4 & \text{if} \quad (x_{t+1}, y_{t+1}) = (x_t, y_t - 1) \\ 1/4 & \text{if} \quad (x_{t+1}, y_{t+1}) = (x_t, y_t + 1) \end{cases}$$

Prove that for $\lambda > 0$

$$Pr\left(|x_t + y_t| \geq \lambda\sqrt{t}\right) \leq 2e^{-\lambda^2/2}.$$

Problem 3 (20 points)

Let $f(X_1, X_2, \ldots, X_n)$ satisfy the Lipschitz condition so that, for any i and any values $x_1, \ldots, x_n$ and y_i,

$$|f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_n)| \leq c.$$

We set

$$Z_0 = \mathbb{E}(f(X_1, \ldots, X_n))$$

and

$$Z_i = \mathbb{E}(f(X_1, \ldots, X_n) \mid X_1, \ldots, X_i).$$

Give an example to show that, if the X_i are not independent, then it is possible that $|Z_i - Z_{i-1}| > c$ for some i.

Problem 4 (20 points)

Consider a bin with $N > 1$ balls. The balls are either black or white. Let $X_0 = \frac{m}{N} < 1$ be the fraction of black balls in the bin at time 0. Let X_i be the fraction of black balls at time i. At step $i \geq 1$ one ball, chosen uniformly at random, is replace with a new ball. With probability X_i the new ball is black, otherwise it is white. All random choices are independent.

Consider the stopping time $\tau := \inf_i\{X_i \in \{0, 1\}\}$. That is, the process stops when all balls have the same color.

(a) Show that $X_1, X_2, \ldots$ is a martingale with respect to itself.

(b) Show that $E[\tau] < \infty$.
Hint: Show that at any step there is probability $\geq (\frac{1}{2N})^{N/2}$ to terminate in the next $N/2$ steps. Conclude that $E[\tau] \leq (2N)^{N/2} + N/2$.

(c) Calculate $\mathbb{P}(X_\tau = 1)$

Problem 5 (20 points)

Consider the range space $([0,1], R)$, where R is the collection of unions of two closed intervals in $[0,1]$. I.e.

$$R = \{[x,y] \cup [v,w] \mid [x,y],\ [v,w] \subseteq [0,1]\}.$$

Assume a distribution $\mathcal{D}$ that is uniform on $[0,1]$.

1. What is the VC-dimension of $([0,1], R)$?
2. Construct an ϵ-net for $([0,1], R)$.
3. Construct an ϵ-sample for $([0,1], R)$. For convenience, you may assume that $\epsilon = \frac{1}{n}$ for a positive integer n.
4. Compare the asymptotic sizes of your constructions in (2) and (3) to the bounds for random sample sizes proven using the VC- dimension. Your constructions should require no more elements than these bounds (up to constant factors).

Problem 6 (20 points)

Consider a range space (X, R), where $X = \{1, 2, \ldots,\}$ (the set of natural numbers) and $R = \{[x, y] \cap X \mid 1 \leq x \leq y \leq \infty\}$. Assume a distribution $\mathcal{D}$ over X defined by $\mathcal{P}(X = i) = 2^{-i}$.

1. What is the VC-dimension of (X, R)?
2. Construct an ϵ-net for (X, R).
3. Construct an ϵ-sample for (X, R).
4. Compare your answers to (2) and (3) to the asymptotic random sample result. Your construction should require no more elements than the random sample results (up to constant factors).