Exercise sheet 2

Exercises for the exercise session on 27 October 2021

Problem 2.1. Let $n \in \mathbb{N}$ and let $\mathcal{F}$ be an inclusion-free family of subsets of $[n] := \{1, 2, \ldots, n\}$ (*inclusion-free* means that no element of $\mathcal{F}$ is a proper subset of another element). Choose a permutation σ of $[n]$ uniformly at random and define the random variable

$$X := |\{k : \{\sigma(1), \sigma(2), \ldots, \sigma(k)\} \in \mathcal{F}\}| .$$

(a) Argue that $\mathbb{E}[X] \le 1$.

(b) Given a fixed set $F \in \mathcal{F}$ of size k, determine the number of permutations σ for which $\{\sigma(1), \sigma(2), \ldots, \sigma(k)\} = F$.

(c) Deduce that

$$\mathbb{E}[X] = \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} .$$

(d) Use (a) and (c) to prove that $|\mathcal{F}| \le \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Problem 2.2. Suppose an unbiased coin is tossed n times. For positive integers $k \le n$, let A_k denote the event that out of these n tosses, there are k consecutive ones with the same outcome (i.e. k consecutive 'heads' or k consecutive 'tails'). Let $\varepsilon > 0$. Prove that

$$\frac{n-k+1}{k 2^{k-1}} \le \mathbb{P}(A_k) \le \frac{n}{2^{k-1}}$$

and deduce that

(a) $\mathbb{P}(A_k) \overset{n\to\infty}{\longrightarrow} 0$ if $k \ge (1+\varepsilon)\log_2 n$;

(b) $\mathbb{P}(A_k) \overset{n\to\infty}{\longrightarrow} 1$ if $k \le \log_2 n - (1+\varepsilon)\log_2 \log_2 n$.

Problem 2.3. Use Markov's inequality to prove Chebyshev's inequality.

Problem 2.4. Let m balls and n bins be given, where m, n are positive integers. Independently for each ball, we choose a bin uniformly at random and place the ball in that bin. For $i = 1, \ldots, n$, denote by A_i the event that the i-th bin remains empty after distributing all balls. The number X of empty bins satisfies $X = X_1 + \cdots + X_n$, where each X_i is the indicator random variable of A_i.

(a) Show that for any constant $\varepsilon > 0$,

$$\mathbb{E}[X] \overset{n\to\infty}{\longrightarrow} \begin{cases} 0 & \text{if } m \ge (1+\varepsilon) n \ln n, \\ \infty & \text{if } m \le (1-\varepsilon) n \ln n. \end{cases}$$

Hint. You can use that $1 + x = \exp\left(x + o\left(x^2\right)\right)$ for $x = o(1)$.

(b) Prove that for any $i \ne j$, X_i, X_j are negatively correlated and deduce that

$$\mathbb{P}(\text{there exists an empty bin}) \overset{n\to\infty}{\longrightarrow} \begin{cases} 0 & \text{if } m \ge (1+\varepsilon) n \ln n, \\ 1 & \text{if } m \le (1-\varepsilon) n \ln n. \end{cases}$$

Problem 2.5. Like in Problem 2.2, let $k \leq n$ be positive integers and toss an unbiased coin n times. For each set S of k consecutive tosses, let A_S be the event that all these tosses have the same outcome. Denote by X_S the indicator random variable of A_S and set $X := \sum_S X_S$. Let $h = h(n)$ be a function that tends to infinity as $n \to \infty$, but satisfies $h = o(\log_2 n)$, and suppose that $k = \log_2(n) - h$.

(a) Show that $\mathbb{E}[X] = \frac{n-k+1}{2^{k-1}}$ and that this tends to ∞.

(b) Suppose that S, T are sets of k consecutive tosses and that $1 \leq |S \cap T| \leq k-1$. Determine $\mathbb{E}[X_S X_T]$.

(c) Prove that

$$\frac{\sum_{S \neq T, X_S, X_T \text{ dependend}} \mathbb{E}[X_S X_T]}{\mathbb{E}[X]^2} \overset{n\to\infty}{\longrightarrow} 0$$

and deduce that $\mathbb{P}(X \geq 1) \overset{n\to\infty}{\longrightarrow} 1$.

Hint. Sort the sets S, T by the value of $j := k - |S \cap T|$.