Probability and Stochastic Processes (2023-24)
Problem Sheet 3

1. Let $X_1, \ldots, X_n$ be 0-1 independent random variables (Poisson trials) such that $\mathbf{P}(X_i) = p_i$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbf{E}[X]$. Then the following Chernoff bounds hold:
   - for any $\delta > 0$, $\mathbf{P}(X \geq (1+\delta)\mu) < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$;
   - for $0 < \delta \leq 1$, $\mathbf{P}(X \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/3}$.
   - for $R \geq 6\mu$, $\mathbf{P}(X \geq R) \leq 2^{-R}$

2. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = \frac{1}{2}$. Let $X = \sum_{i=1}^n X_i$. For any $a > 0$, $\mathbf{P}(X \geq a) \leq e^{-a^2/2n}$. Also, $\mathbf{P}(X \leq -a) \leq e^{-a^2/2n}$.

3. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = \frac{1}{2}$. Let $X = \sum_{i=1}^n X_i$. For any $a > 0$, $\mathbf{P}(|X| \geq a) \leq 2e^{-a^2/2n}$.

4. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = 0) = \frac{1}{2}$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbf{E}[X] = \frac{n}{2}$.
   - for any $a > 0$, $\mathbf{P}(X \geq \mu + a) \leq e^{-2a^2/n}$.
   - for any $\delta > 0$, $\mathbf{P}(X \geq (1+\delta)\mu) \leq e^{-\delta^2\mu}$.

5. Let $X_1, \ldots, X_n$ be independent random variables such that $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = 0) = \frac{1}{2}$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbf{E}[X] = \frac{n}{2}$.
   - for any $0 < a < \mu$, $\mathbf{P}(X \leq \mu - a) \leq e^{-2a^2/n}$.
   - for any $0 < \delta < 1$, $\mathbf{P}(X \leq (1-\delta)\mu) \leq e^{-\delta^2\mu}$.

6. Suppose we roll a standard fair die 200 times. Let X be the sum of the numbers that appear over the 200 rolls. Use Chebyshev's inequality to bound $\mathbf{P}[X \geq 750]$.

7. Let X be the random variable denoting the number of heads in a sequence of n independent fair coin flips. If the following Chernoff bound $\mathbf{P}\left(|X - \mu| \geq \delta\mu\right) \leq 2e^{-\mu\delta^2/3}$ is applied to get

$$\mathbf{P}\left(\left|X - \frac{n}{2}\right| \geq Y\right) \leq \frac{2}{3},$$

   then what is Y?

8. Identical jobs are assigned to identical machines uniformly and independently at random. Show that when there are n processors and $\Omega(n \log n)$ jobs, then with high probability, every processor will have a load between half and twice the average.

9. A tournament on a set V of n players is an orientation $T = (V, E)$ of the edges of the complete graph on the set of vertices V. Thus, for every two distinct elements x, y of V either (x, y) or (y, x) is in E, but not both. We say that T has the property S_k if for every set of k players there is one player who beats all k of them. For example, a directed triangle $T_3 = (V, E)$, where $V = \{1, 2, 3\}$ $E = \{(1, 2), (2, 3), (3, 1)\}$, has S_1.

   If $\binom{n}{k}\left(1 - 2^{-k}\right)^{n-k} < 1$ then there is a tournament on n vertices that has the property S_k.

10. There is a tournament T with n players and at least $\frac{n!}{2^{(n-1)}}$ Hamiltonian paths.

11. Given an undirected graph G with m edges, there is a partition of V into two disjoint sets A, B such that at least $m/2$ edges connect A to B, that is, there is a cut of value at least $m/2$.

12. There is a two-coloring of edges of K_n with at most

$$\binom{n}{a} 2^{1-\binom{a}{2}}$$

   monochromatic K_a.

13. For a set S of n points in the unit square U, let $T(S)$ be the area of the smallest triangle whose vertices are three distinct points of S. There is a set S of n points in the unit square U such that $T(S) \geq \frac{1}{100n^2}$.

14. Suppose you are given a biased coin with bias $p > 1/2$. Can you use it to simulate an unbiased coin toss? Explain.

    Also, compute the number of tosses you would need on expectation to simulate an unbiased coin toss.

15. Real-life computers only have access to uniform random number generators. Given any distribution $\mathcal{D}$, provide an idea of how to generate samples from $\mathcal{D}$, assuming you have access to a uniform generator only.

16. Are the Markov and Chebyshev's inequalities tight? If yes, give examples. If not, give explanations as to why.

17. Given a random variable X defined as $X = \sum_{i=1}^{n} X_i$ where $\mathbf{E}[X_i X_j] = \mathbf{E}[X_i]\mathbf{E}[X_j]$ for all $i \neq j$, prove that $Var[X] = \sum_{i=1}^{n} Var[X_i]$.

18. For a random variable X with standard deviation $\sigma(X)$ and a positive real t, show that:

$$\Pr(X - E[X] \geq t\sigma(X)) \leq \frac{1}{1 + t^2}$$

    Derive a two sided Chebyshev's like bound from this result. Find out if there exists some cases where it would give a better bound than the standard Chebyshev's inequality.

19. For a non-negative integer valued random variable X, show that:

$$\frac{E[X]^2}{E[X^2]} \leq \Pr(X \neq 0) \leq E[X]$$

20. (a) Determine the moment-generating function for the binomial random variable $B(n, p)$.

    (b) Let X be a $B(n, p)$ random variable and Y a $B(m, p)$ random variable, where X and Y are independent. Use part (a) to determine the moment-generating function of $X + Y$.

    (c) What can we conclude from the form of the moment-generating function of $X + Y$?