DCP1206 Fall 2019: Probability **(Due: 2019/12/11 in class)**

Homework 5: MGF and Concentration Inequalities

Problem 1 (Moment Generating Functions) (10+10=20 points)

Let X be a random variable with the PDF given by

$$f_X(x) = \begin{cases} \frac{1}{4}, & \text{if } x \in (-1,3) \\ 0, & \text{else} \end{cases}$$

(a) Find $M_X(t)$, i.e. the moment generating function of X.

(b) Use $M_X(t)$ to find $E[X]$ and $\mathrm{Var}[X]$.

Problem 2 (Use MGF to find distributions) (6+6+6=18 points)

In each of the following cases, $M_X(t)$, the moment generating functions of X, is given. Please determine the distribution of X. (You could use the MGF table in the slides or the one in the textbook)

(a) $M_X(t) = \left(\frac{1}{4}e^t + \frac{3}{4}\right)^7$.

(b) $M_X(t) = e^t/(2-e^t)$.

(c) $M_X(t) = \exp\left[3(e^t-1)\right]$.

Problem 3 (Sum of independent random variables) (12+12=24 points)

(a) Let X_1 and X_2 be two independent continuous uniform random variables on $(0,2)$. Use convolution theorem to find the PDF of $X_1 + X_2$.

(b) Let $Y_1 \sim \mathcal{N}(1,4)$ and $Y_2 \sim \mathcal{N}(4,9)$ be independent random variables. Find the probability of the event $3Y_1 + 4Y_2 > 20$.

Problem 4 (Concentration Inequalities) (10+14=24 points)

(a) Suppose that X is a random variable with $E[X] = \mathrm{Var}[X] = \mu$. What does Chebyshev's inequality say about $P(X > \pi\mu)$?

(b) Imagine we have an algorithm for solving some decision problem (e.g., is a given number a prime?). Suppose that the algorithm makes a decision at random and returns the correct answer with probability $\frac{1}{2}+\delta$, for some $\delta > 0$ (which is just a bit better than a random guess). To improve the performance, we run the algorithm N times and take the majority vote. Show that, for any $\epsilon \in (0,1)$, the answer is correct with probability at least $1-\varepsilon$, as long as $N \geq (1/2)\delta^{-2}\ln(\varepsilon^{-1})$. (Hint: For each i, define X_i to be a Bernoulli random variable for which $X_i = 1$ when the algorithm return the correct answer at the i-th trial. Under majority vote, we know that if the final answer is incorrect, then $X_1 + X_2 + \cdots + X_N \leq N/2$. Use the negative part of Hoeffding's inequality. This scheme is usually called "boosting randomized algorithms.")

Problem 5 (Techniques of Chernoff Bound) (10+14=24 points)

Let $X_1, \cdots, X_N$ be non-negative independent random variables with continuous distributions (but $X_1, \cdots, X_N$ are not necessarily identically distributed). Assume that the PDFs of X_i's are uniformly bounded by 1.

(a) Show that for every i, $E[\exp(-tX_i)] \leq \frac{1}{t}$, for all $t > 0$.

(b) By using (a), show that for any $\varepsilon > 0$, we have

$$P\Big(\sum_{i=1}^{N} X_i \leq \varepsilon N\Big) \leq (e\varepsilon)^N.$$

(Hint: For any $t > 0$, $P(\sum_{i=1}^N X_i \leq \varepsilon N) = P(e^{t\sum_{i=1}^N X_i} \leq e^{t\varepsilon N}) = P\big(e^{-t\sum_{i=1}^N X_i} \geq e^{-t\varepsilon N}\big)$)