CS5314 Randomized Algorithms

Homework 3 Suggested Solution

(Homework due date was May 19, 2020)

1. (a) Determine the moment generating function for the binomial random variable $Bin(n, p)$.
   Ans. Let $X \sim Bin(n, p)$. Let $X_1, X_2, \ldots, X_n$ be independent indicators, each with probability p to be successful. Then, $X = \sum_i X_i$ and MGF for X is the product of the MGFs for X_i.
   For any X_i, its MGF is:
   $$\mathrm{E}[e^{tX_i}] = p \cdot e^t + (1-p).$$
   Thus, the MGF for X is:
   $$\prod_i \mathrm{E}[e^{tX_i}] = (p\, e^t + (1-p))^n.$$

   (b) Let X be a $Bin(n, p)$ random variable and Y be a $Bin(m, p)$ random variable. Suppose that X and Y are independent. Use part (a) to determine the moment generating function of $X + Y$.
   Ans. Since X and Y are independent, the MGF for $X + Y$ is equal to the product of the MGF for X and the MGF for Y.
   Based on the result in part (a), the MGF for $X + Y$ is:
   $$\mathrm{E}[e^{tX}] \cdot \mathrm{E}[e^{tY}] = (p \cdot e^t + (1-p))^n \; (p \cdot e^t + (1-p))^m = (p \cdot e^t + (1-p))^{n+m}.$$

   (c) What can we conclude from the form of the moment generating function of $X + Y$?
   Ans. The MGF for $X + Y$ is of the same form as the MGF for $Bin(n + m, p)$. This implies that $X + Y$ has the same distribution as $Bin(n + m, p)$. In other words, sum of independent binomial random variables with the same parameter p is a binomial random variable.

2. Let $X_1, X_2, \ldots, X_n$ be independent Poisson trials such that $\Pr(X_i) = p_i$. Let $X = \sum_{i=1}^n X_i$ and $\mu = \mathrm{E}[X]$. During the class, we have learnt that for any $\delta > 0$,
   $$\Pr(X \geq (1+\delta)\mu) < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu.$$
   In fact, the above inequality holds for the weighted sum of Poisson trials. Precisely, let $a_1, \ldots, a_n$ be real numbers in $[0, 1]$. Let $W = \sum_{i=1}^n a_i X_i$, and $\nu = \mathrm{E}[W]$. Then, for any $\delta > 0$,
   $$\Pr(W \geq (1+\delta)\nu) < \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\nu.$$

   (a) Show that the above bound is correct.
   Ans. Since $W = \sum_{i=1}^n a_i X_i$, we have
   $$\nu = \mathrm{E}[W] = \sum_{i=1}^n a_i \mathrm{E}[X_i] = \sum_{i=1}^n a_i p_i$$

For any i,

$$\mathrm{E}[e^{ta_iX_i}] = p_i e^{ta_i} + (1-p_i) = 1 + p_i(e^{ta_i} - 1) \le e^{p_i(e^{ta_i}-1)}$$

Claim 1. For any $x \in [0,1]$, $e^{tx} - 1 \le x(e^t - 1)$
Proof.

$$\begin{aligned} & f(x) = x(e^t-1) - e^{tx} + 1 \\ \Rightarrow\ & f'(x) = (e^t - 1) - te^{tx} \\ \Rightarrow\ & f'(x) = 0 \text{ (when } x = x^* = (\ln(e^t-1) - \ln t)/t) \\ \Rightarrow\ & f''(x) = -t^2 e^{tx} \le 0 \end{aligned}$$

In other words, for $x \in [0,1]$, $f(x)$ achieves minimum value either at $f(0)$ or $f(1)$. So $f(x) \ge \min\{f(0), f(1)\} = 0$ for all $x \in [0,1]$. □

Hence,

$$\mathrm{E}[e^{eta_iX_i}] \le e^{p_i(e^{ta_i}-1)} \le e^{a_ip_i(e^t-1)}$$

By the independence of X_i's and the property of MGF,

$$\mathrm{E}[e^{tW}] = \prod_{i=1}^{n} \mathrm{E}[e^{ta_iX_i}] \le \prod_{i=1}^{n} \mathrm{E}[e^{a_ip_i(e^t-1)}] = e^{\nu(e^t-1)}$$

For any $t > 0$, we have

$$\Pr(W \ge (1+\delta)\nu) = \Pr(e^{tW} \ge e^{t(1+\delta)\nu}) \le \frac{\mathrm{E}[e^{tW}]}{e^{t(1+\delta)\nu}} \le \frac{e^{\nu(e^t-1)}}{e^{t(1+\delta)\nu}}$$

Then, for any $\delta > 0$, we can set $t = \ln(1+\delta) > 0$ and obtain:

$$\Pr(W \ge (1+\delta)\nu) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\nu}$$

(b) Prove a similar bound for the probability $\Pr(W \le (1-\delta)\nu)$ for any $0 < \delta < 1$.
Ans. For any $t < 0$, we have

$$\Pr(W \le (1-\delta)\nu) = \Pr(e^{tW} \ge e^{t(1-\delta)\nu}) \le \frac{\mathrm{E}[e^{tW}]}{e^{t(1-\delta)\nu}} \le \frac{e^{\nu(e^t-1)}}{e^{t(1-\delta)\nu}}$$

Then, for any $0 < \delta < 1$, we can set $t = \ln(1-\delta) < 0$ and obtain:

$$\Pr(W \le (1-\delta)\nu) < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\nu}$$

3. Let $X_1, \ldots, X_n$ be independent random variables such that

$$\Pr(X_i = 1 - p_i) = p_i \quad \text{and} \quad \Pr(X_i = -p_i) = 1 - p_i.$$

Let $X = \sum_{i=1}^{n} X_i$. Prove that

$$\Pr(|X| \ge a) \le 2e^{-2a^2/n}.$$

Note: You may assume that the following inequality, which is a special case of **Hoeffding's Lemma**, is correct:

$$p_i e^{\lambda(1-p_i)} + (1-p_i)e^{-\lambda p_i} \le e^{\lambda^2/8}.$$

Ans. For each X_i, its MGF is of the form:

$$\mathrm{E}[e^{tX_i}] = p_i e^{t(1-p_i)} + (1-p_i)e^{-tp_i}$$

By Hoeffding's Lemma, we have:

$$\mathrm{E}[e^{tX_i}] = p_i e^{t(1-p_i)} + (1-p_i)e^{-tp_i} \le e^{t^2/8}$$

Thus, by the independence of X_is and the property of MGF,

$$\mathrm{E}[e^{tX}] \le e^{nt^2/8}$$

Hence, for $t = 4a/n > 0$, we have:

$$\Pr(X \ge a) \le \Pr(e^{tX}/e^{ta}) \le \mathrm{E}[e^{tX}]/e^{ta} \le e^{nt^2/8}/e^{ta} = e^{-2a^2/n}$$

Also, for $t = -4a/n < 0$, we have:

$$\Pr(X \ge -a) \le \Pr(e^{tX}/e^{-ta}) \le \mathrm{E}[e^{tX}]/e^{-ta} \le e^{nt^2/8}/e^{-ta} = e^{-2a^2/n}$$

Combining the above two bounds yield the desired bound.

4. (No marks) Study Hoeffding's Lemma.[1]

[1]Check this out: `https://en.wikipedia.org/wiki/Hoeffding%27s%5flemma`