Homework 2
(150 points)

1. **(10 + 10 + 20 points) An Interesting Concentration.** Let $\mathbb{X}$ be the random variable over the sample space $\{1, 2, \dots\}$ such that $\mathbb{P}[\mathbb{X} = i] = 2^{-i}$.

   (a) Compute $\mu = \mathbb{E}[\mathbb{X}]$.

   (b) Define $\mathbb{Y} = \mathbb{X} - \mu$. For $0 \leqslant h \leqslant \ln 2$, compute $\mathbb{E}\left[\exp(h\mathbb{Y})\right]$.

   (c) Define $\mathbb{S}_n = \mathbb{Y}^{(1)} + \cdots + \mathbb{Y}^{(n)}$. Find the concentration bound for $\mathbb{P}[\mathbb{S}_n \geqslant t]$ using the technique of Chernoff bound.

2. **(10 + 10 + 20 points) Concentration of Sum of Poisson Distribution.** Let $\mathbb{X}$ be the random variable over the sample space $\{0, 1, \dots\}$ such that $\mathbb{P}[\mathbb{X} = i] = \exp(-\mu)\frac{\mu^i}{i!}$.

   (a) Prove that $\mathbb{E}[\mathbb{X}] = \mu$.

   (b) Define $\mathbb{Y} = \mathbb{X} - \mu$. For positive h, compute $\mathbb{E}\left[\exp(h\mathbb{Y})\right]$.

   (c) Define $\mathbb{S}_n = \mathbb{Y}^{(1)} + \cdots + \mathbb{Y}^{(n)}$. Find the concentration bound for $\mathbb{P}[\mathbb{S}_n \geqslant t]$ using the technique of Chernoff bound. (You might find it useful to use a variable m such that $m = n\mu$ in the final bound.)

3. **(10 + 10 points) Coin Tossing.** Let $\mathbb{X}$ be the uniform distribution over the sample space $\{0, 1\}$.

   (a) Let $S_n = \mathbb{X}^{(1)} + \cdots + \mathbb{X}^{(n)}$. Given a fixed values of m, how will you choose n such that $\mathbb{P}\left[\mathbb{S}_n \geqslant m\right] \leqslant (1 - \varepsilon)$?

   (b) Use the above result to prove the concentration bound in Problem 1 part c.

4. **(10 points) Concentration of Matrix rank.** Let $\mathbb{M}$ be a distribution over $n \times n$ matrices, where each element is selected uniformly and independently at random from the set Ω. State and prove a concentration bound for the rank of $\mathbb{M}$ around its median or mean.

5. **(40 points) Prefix-sum of Coins are Close to their respective Mean.** Let $\mathbb{X}$ be a distribution over $\{0,1\}$ such that $\mathbb{P}[\mathbb{X}=1]=p$ and $\mathbb{P}[\mathbb{X}=0]=(1-p)$. We consider the sum $S_n = \mathbb{X}^{(1)} + \cdots + \mathbb{X}^{(n)}$.

Chernoff-Hoeffding's bound states the following. It says that the probability of the sum S_n exceeding the expectation by t is very small. For example, we can say that

$$\mathbb{P}[\mathbb{S}_n \geqslant pn + t] \leqslant \exp\left(-2t^2/n\right)$$

Intuitively, suppose we reject any outcome of the coins such that $\mathbb{S}_n \geqslant pn + t$. Then, this bound says that the probability of rejecting is at most $\exp\left(-2t^2/n\right)$.

We want to claim that "$\mathbb{S}_n$ never exceeded the expectation in any prefix." Let me elaborate. Suppose we reject any coin such that $\mathbb{S}_i \geqslant p \cdot i + t$ for any $i \in \{1,\ldots,n\}$. Formally, we reject if there exists $i \in \{1,\ldots,n\}$ such that $\mathbb{S}_i \geqslant p \cdot i + t$. Note that this rejection rule is *more stringent* than the previous rejection criterion. Our goal is to prove that this rejection probability is small. In particular, prove that

$$\mathbb{P}\left[\exists i \in \{1,\ldots,n\} \text{ s.t. } \mathbb{S}_i \geqslant p \cdot i + t\right] \leqslant \exp\left(-2t^2/n\right)$$

Isn't this amazing? This bound is identical to the Chernoff-Hoeffding bound!

6. **(Extra Credit) New bounds for Hoeffding's Lemma.** Surprise me with a new statement/proof of Hoeffding's Lemma!