Collaboration Policy: You may collaborate with other students on these problems. Collaboration is limited to discussion of ideas only, and you should write up the solutions entirely on your own and list your collaborators as well as cite any references you may have used.

1. **Definition 0.1** A family of hash functions $H = \{h : M \rightarrow N\}$ is said to be a perfect hash family if for each set $S \subseteq M$ of size $s \leq n$, there exists a hash function $h \in H$ that is perfect for $S$.

   Assuming that $n = s$, show that any perfect hash family must have size $2^{\Omega(s)}$.

2. Consider the following generalization of perfect hash functions.

   **Definition 0.2** Let $b_i(h, S) = \{|x \in S| h(x) = i\}$. A hash function $h$ is $b$-perfect for $S$ if $b_i(h, S) \leq b$ for each $i$. A family of hash functions $H = \{h : M \rightarrow N\}$ is said to be a $b$-perfect hash family if for each $S \subseteq M$ of size $s$ there exists a hash function $h \in H$ that is $b$-perfect for $S$.

   Show that there exists a $b$-perfect hash family $H$ such that $b = O(\log n)$ and $|H| \leq m$, for any $m \geq n$.

3. We analyzed the process of throwing $n$ balls into $n$ bins independently and at random and showed that the maximum load is at most $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

   (a) Suppose instead that the balls are assigned to bins by a function $f$ which is chosen from a 2-universal family. Establish an upper bound on the maximum load that holds with probability at least $1/2$ in this case.

   (b) Suppose we have a family of hash functions $\mathcal{H} = \{h : [n] \rightarrow [n]\}$ such that, for all $x_1, \ldots, x_k, y_1, \ldots, y_k \in [n],$

   $$\Pr_{h \in \mathcal{H}}[(h(x_1) = y_1) \land (h(x_2) = y_2) \land \ldots \land (h(x_k) = y_k)] = O\left(\frac{1}{n^k}\right)$$

   Suppose the allocation of balls into bins is done using a function $h \in \mathcal{H}$ chosen at random. Obtain an upper bound on the maximum load that holds with probability at least $1/2$ in this case.
4. Prove the following extension of the Chernoff bound we proved in class. Let $X = \sum_{i=1}^{n} X_i$ where the $X_i$ are independent $0-1$ random variables. Let $\mu = E[X]$. Suppose that $\mu_L \leq \mu \leq \mu_H$. Then for any $\delta > 0$,

$$\Pr(X \geq (1 + \delta)\mu_H) \leq \left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}}\right)^{\mu_H}.$$ 

For $0 < \delta < 1$,

$$\Pr(X \leq (1 - \delta)\mu_L) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}}\right)^{\mu_L}.$$