1. **Probability Review:**

   (a) Imagine you are given one fair die, and you need to decide which task is harder: (i) guessing the value of one die toss or (ii) tossing the die twice and getting the same value twice. Given that the die is fair (every side has weight 1/6), does event (i) have a greater chance of success or event (ii), or do they have the same probability of success? Make sure to give justification.

   (b) We will now generalize this result to $n$-sided dice with any (possibly non-uniform) distribution. First prove the following useful fact, for any $\alpha_1, \alpha_2, \ldots, \alpha_n$ such that $\sum \alpha_i = 1$, the following holds,

   $$0 \leq \sum_{i=1}^{n} (\alpha_i - 1/n)^2 = \sum_{i=1}^{n} \alpha_i^2 - 1/n$$

   (c) Let $X_1$ be the value of the first toss, and $X_2$ be the value of the second toss. Show that $\Pr(X_1 = X_2) \geq 1/n$ (hint: use part b). For what distribution is the inequality tight?

2. **Concentration Bounds:**

   (a) Given a sample of $m$ bounded points $X = (x_1, x_2, \ldots, x_m)$, $\forall i, |x_i| \leq M$, define the function

   $$f(X) = \frac{1}{m} \sum_{i} x_i,$$

   Can you give a bound on the probability $\Pr[|f(X) - \mathbb{E}[f(X)]| \geq \epsilon]$?

   (b) Let $X$ and $X'$ be two sets of size $m$ that differ in exactly one point. That is, $|X \cap X'| = m - 1$. We say a function $h$ is *stable* if for all such $X, X'$, $|h(X) - h(X')| \leq g(m)$ for some decreasing function $g$. How quickly does $g$ need to decrease as a function of
In order for McDiarmid’s inequality to provide a bound on the event \( \Pr[|h(X) - E[h(X)]| \geq \epsilon] \) that converges to zero as \( m \to \infty \)?

(c) Is the function \( f \) from part (a) stable (still assuming the bound \( |x_i| \leq M, \forall i \))? Will McDiarmid’s inequality provide provide a convergent bound? If so give the bound. Now define the function \( f'(X) = \max(X) \), is \( f' \) stable? Can you give a bound with McDiarmid’s inequality?

3. PAC Learning: Here we will consider an alternative PAC learning scenario, called the two-oracle model. Imagine you are given the ability to explicitly ask for a positive or negative sample, which are drawn from different distributions \( D_+ \) and \( D_- \) respectively. A concept is efficiently PAC-learnable if there exists an algorithm \( L \) that can generate a hypothesis \( h \), such that \( \Pr_{x \sim D_+}[h(x) = 0] \leq \epsilon \) and \( \Pr_{x \sim D_-}[h(x) = 1] \leq \epsilon \) with confidence \((1 - \delta)\), after sampling \( m = \text{poly}(1/\epsilon, 1/\delta) \) points.

(a) Show that if a problem is efficiently PAC-learnable in the classic sense, it is also always efficiently PAC-learnable in the two-oracle model.

(b) (Bonus) Show that the reverse direction is also true.