Question 1

In Lecture 3 we showed that the inner product function $IP(x, y) = x_1y_1 + \cdots + x_n y_n \mod 2$, where $x, y \in \{0, 1\}^n$ takes the same value on more than $7/8$ of the entries of any set of the form $X \times Y$ where $|X| \cdot |Y| \geq K \cdot 2^n$ for some constant $K$. In this question you will show that the same is true with high probability for a random function $R: \{1, \ldots, N\} \times \{1, \ldots, N\} \to \{0, 1\}$, where $N = 2^n$.

(a) Let $Z_1, \ldots, Z_M$ be a sequence of independent uniformly random coin tosses. Apply the inequality $\binom{M}{\delta M} \leq 2^{H(\delta)M}$ and a union bound to show that the probability more than $7M/8$ of the coins are heads is at most $2^{-M/4}$.

**Solution:** The probability that any $7M/8$ specific $Z_i$’s are heads is $2^{-7M/8}$. By a union bound, the probability that there exists some set of $7M/8$ heads is at most $\binom{M}{7M/8} \cdot 2^{-7M/8} \leq 2^{H(7/8) - 7/8)M}$ which is at most $2^{-M/4}$ as $H(7/8) - 7/8 \geq -0.33$.

(b) Use part (a) to show that the probability $R$ takes the same value on more than $7/8$ of the entries of some set of the form $X \times Y$ is at most $2^{-|X| \cdot |Y|/4+1}$.

**Solution:** The values $R(x, y)$ where $x \in X$ and $y \in Y$ are $|X| \cdot |Y|$ independent bits. By part (a) the probability that a $7/8$ fraction of them are zeros is at most $2^{-|X| \cdot |Y|/4}$. The same bound holds for ones. By a union bound the probability that a $7/8$ fraction of values are equal is at most $2^{-|X| \cdot |Y|/4+1}$.

(c) Use part (b) and a union bound to show that a random function takes the same value on more than $7/8$ of the entries of some set $X \times Y$ with $|X| \cdot |Y| \geq 9N$ with probability at most $2^{-\Omega(N)}$.

**Solution:** By part (b), assuming $|X| \cdot |Y| \geq 9N$, the probability of the event is at most $2^{-9N/4+1}$. There are at most $2^{2N}$ pairs of subsets $X, Y$. By a union bound the probability that there exists a subset that has the property is at most $2^{2N} \cdot 2^{-9N/4+1} = 2^{-N/4+1} = 2^{-\Omega(N)}$.

Question 2

Given an undirected graph $G$, let $G^2$ be the graph whose vertices are ordered pairs of vertices in $G$ and whose edges are those pairs $\{(u, v), (u', v')\}$ such that $\{u, u'\}$ is an edge in $G$ or $u = u'$, and $\{v, v'\}$ is an edge in $G$ or $v = v'$.

(a) Show that if $G$ has a clique of size $k$ then $G^2$ has a clique of size $k^2$.

**Solution:** If $S$ is the set of $k$ vertices in $G$ that forms a clique, then $S^2 = \{(u, v) : u, v \in K\}$ is a set of $k^2$ vertices that is a clique in $G^2$.

(b) Show that if $G^2$ has a clique of size $K$ then $G$ has a clique of size $\lceil \sqrt{K} \rceil$.

**Solution:** Let $T$ be a clique in $G^2$ and $U = \{u : (u, v) \in T\}$, $V = \{v : (u, v) \in T\}$ be its projections to vertices in $G$. Then $U$ and $V$ are clique in $G$: If $\{(u, v), (u', v')\}$ is an edge in $G^2$ then by the definition of $G^2$ both $(u, u')$ and $(v, v')$ must be edges in $G$. Since $T$ is contained in the set $U \times V$, it follows that $|U| \cdot |V| = |U \times V| \geq |T|$. If $T$ has size $K$ then either $U$ or $V$ must then have size at least $\lceil \sqrt{K} \rceil$ as desired.
(c) Use parts (a) and (b) to show that if there exists a polynomial-time algorithm that finds a clique of size at least 1% of the size of the largest clique in a graph, then there is a polynomial-time algorithm that finds a clique of size at least 99% the size of the largest clique.

**Solution:** Let $A$ be an algorithm that finds a clique of size $\delta$ times the size of the largest clique. The reduction $R$ runs $A$ on the graph $G^2$ to obtain a clique $T$ and outputs the larger of the two sets $U$ and $V$ from part (b). This is a polynomial-time algorithm. By part (a), if $G$ has a clique of size $k$ then $G^2$ has a clique of size $k^2$. By our assumption on $A$, $T$ is then a clique of size at least $\delta \cdot k^2$. By part (b), the reduction outputs a clique in $G$ of size at least $\sqrt{\delta k^2} = \delta^{1/2} \cdot k$.

Composing $R$ with itself 9 times, we obtain a polynomial-time reduction from finding a clique of size $\delta^{1/2^9}$-fraction of the largest one to finding one of size $\delta$-fraction of the largest one. When $\delta = 1\%$, $\delta^{1/2^9} \geq 99\%$ as desired.

**Question 3**

A function $f : \{0,1\}^n \to \{0,1\}$ is affine if it is of the form $f(x) = \langle a, x \rangle + b$ for some $a \in \{0,1\}^n$ and $b \in \{0,1\}$. It is $\delta$-far from affine if every affine function differs from it on more than a $\delta$-fraction of inputs. The YES and NO instances of $(1, 1-\delta)$-GAP-AFFINE are functions that are affine and $\delta$-far from affine, respectively.

(a) Let $g(x, y) = f(x) + f(y)$. Show that if $f$ is affine then $g$ is linear.

**Solution:** $g(x, y) = (\langle a, x \rangle + b) + (\langle a, y \rangle + b) = \langle a, x \rangle + \langle a, y \rangle = \langle a, a \rangle, (x, y)$.

(b) Show that if $g$ is $\delta$-close to linear then $f$ is $\delta$-close to affine. (Hint: Fix $y$.)

**Solution:** If $\Pr[g(x, y) = \langle a, x \rangle + \langle b, y \rangle] \leq \delta$, then the same inequality must hold for some fixing of $y = c$ that minimizes the left-hand side. It follows that $\Pr[f(x) + f(c) = \langle a, x \rangle + \langle b, c \rangle] \geq \delta$, so $f(x)$ is $\delta$-close to the affine function $\langle a, x \rangle + (\langle b, c \rangle + f(c))$.

(c) Use part (a) and results from Lecture 11 to show that the one-sided randomized query complexity of $(1, 1-\delta)$-GAP-AFFINE with error $1-\delta$ is at most 6.

**Solution:** The test chooses random inputs $x, y, x', y'$ and accepts if $f(x) + f(y) + f(x') + f(y') = f(x + x') + f(y + y')$. If $f$ is affine then by part (a) $g$ is linear and the test accepts with probability 1. By Claim 9 in Lecture 10, if the test accepts with probability $1 - \delta$ then $g$ is $\delta$-close to linear. By part (b) $f$ is then $\delta$-close to affine.

(d) Show that for every three distinct points $x, y, z \in \{0,1\}^n$ and values $a, b, c \in \{0,1\}$ there exists an affine function $f$ such that $f(x) = a$, $f(y) = b$, and $f(z) = c$.

**Solution:** First we argue that there is always a linear function consistent with two constraints $f(x) = a$, $f(y) = b$ where $x, y$ are distinct and nonzero. There is always some index $i$ for which $x_i \neq y_i$. Without loss of generality assume $x_i = 1$ and $y_i = 0$. Let $y_j$ be any 1-input of $y$ and $s \in \{0,1\}^n$ be a string with $s_j = b, s_i = a + bx_j,$ and zero everywhere else. Then $\langle s, x \rangle = (a + bx_j)x_i + bx_j = a$ and $\langle s, y \rangle = s_jy_j = b$ so the linear function $f(u) = \langle s, u \rangle$ satisfies both constraints.

For the problem at hand let $g(u) = f(u + x) + a$. By what we just proved there is a linear function $\langle s, u \rangle$ such that $g(y + x) = ips, y + x$ and $g(z + x) = \langle s, z + x \rangle$. Then the affine function $\langle s, u \rangle + \langle (s, x) + a \rangle$ satisfies all three constraints $f(x) = a$, $f(y) = b$, and $f(z) = c$.  
(e) Use part (b) to show that the one-sided randomized query complexity of $(1, 1 - \delta)$-GAP-AFFINE with any error less than one is at least 4.

**Solution:** Suppose there is an algorithm with query complexity 3 (or less). After querying any three values $x, y, z$ and receiving answers $a, b, c$ a one-sided test must accept because there is at least one function $f$ such that $f(x) = a$, $f(y) = b$, and $f(z) = c$. Therefore the test accepts all functions, including the ones that are $\delta$-far from affine.