Problem 1:
This is a quick refresher for basic probability concepts. A probability distribution over a (finite) set $S$ is a function $\pi : S \to [0, 1]$ such that $\sum_{x \in S} \pi(x) = 1$. A (real-valued) random variable $X$ is a function $X : S \to \mathbb{R}$ along with a probability distribution $\pi$. We define $\Pr_{x \sim \pi}[X(s) = x] = \sum_{x \in S} \pi(x)$ (often shortened to $\Pr[X = x]$, when $\pi$ is understood). We define expectation $\mathbb{E}_{x \sim \pi}[X(s)] = \sum_{s \in S} X(s) \cdot \pi(s)$ (often shortened to $\mathbb{E}[X]$, when $\pi$ is understood).

(a) (Linearity of expectation.) Given two random variables $X_1, X_2$, define a new random variable $X$ as $X(s) = aX_1(s) + bX_2(s)$ (for some real numbers $a$ and $b$). Show that $\mathbb{E}[X(s)] = a\mathbb{E}[X_1(s)] + b\mathbb{E}[X_2(s)]$.

(b) (Markov’s inequality.) Given a non-negative random variable $X$, show that $\Pr[X > t\mu] < 1/t$, where $\mu = \mathbb{E}[X]$.

(c) Given a random variable $X$, suppose we define a new random variable $Z_X$ as $Z_X(s) = X(s) - \mu$ where $\mu = \mathbb{E}[X]$. Calculate $\mathbb{E}[Z_X]$.

(d) (Variance and Chebyshev’s inequality.) Given a random variable $X$, define a new random variable $Z_X$ as $Z_X(s) = (X(s) - \mu)^2$ where $\mu = \mathbb{E}[X]$. Then the variance of $X$ is defined as $\text{Var}(X) = \mathbb{E}[Z_X]$ and the standard deviation as $\sigma(X) = \sqrt{\mathbb{E}[Z_X]}$. Use Markov’s inequality to bound $\Pr[X - \mu > t\sigma(X)]$.

(e) Two random variables $X$ and $Y$ are said to be independent if for all real numbers $x, y$, $\Pr[X = x \land Y = y] = \Pr[X = x] \cdot \Pr[Y = y]$. Show that if $X$ and $Y$ are independent, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. Further, if $\{X_i\}_{i=1}^t$ are $t$ random variables which are pairwise independent (that is, $X_i$ and $X_j$ are independent for all $i \neq j$), show that $\text{Var}(\sum_i X_i) = \sum_i \text{Var}(X_i)$.

(f) Suppose $\{X_i\}_{i=1}^t$ are $t$ pairwise independent random variables which take binary (0-1) values such that $\Pr[X_i = 1] = p$ for all $i$. Use Chebyshev’s inequality to prove that

$$\Pr\left[\left|\frac{\sum_{i=1}^t X_i}{t} - p\right| > \delta\right] = O\left(\frac{1}{\delta^2 t}\right).$$

Problem 2:
Let $M$ be a probabilistic TM. Define the gap of $M$ for a language $L$ to be $\min_{x \in L} \Pr[M(x) = \text{yes}] - \max_{x \notin L} \Pr[M(x) = \text{yes}]$, and its error for $L$ to be $\max_x \Pr[M(x) \neq L(x)]$. Bound the gap and error in terms of each other.

Problem 3:
Define Expected-Time-$\text{PP}$ to be the class of languages decided by probabilistic Turing machines (via acceptance probability $> \frac{1}{2}$) whose expected running-time is polynomial (as opposed to $\text{PP}$, where the running time is worst-case polynomial). Show that $\text{EXP} \subseteq \text{Expected-Time-PP}$. What can you say about inclusion in Expected-Time-$\text{PP}$ for classes larger than $\text{EXP}$? What if the expected running time is restricted to be constant instead of polynomial?

Problem 4:
In this problem we shall prove impossibility of deterministic extraction from Santha-Vazirani sources. We work with probability distributions over $S = \{0, 1\}^n$, the set of $n$-bit strings.

For $x \in \{0, 1\}^n$, let $x_i$ denote the $i$-th bit of $x$ and $x^-i$ denote the other $n-1$ bits of $x$. Call a distribution $\pi$ $\delta$-balanced at position $i$ if for all $y \in \{0, 1\}^{n-1}$, $\Pr[x_i = 0 | x^-i = y]$ and $\Pr[x_i = 1 | x^-i = y]$ differ by at most $\delta$.

(a) Verify that $\pi$ is $\delta$-balanced at position $i$ if and only if for every $y \in \{0, 1\}^{n-1}$,

$$\frac{1 - \delta}{1 + \delta} \leq \frac{\pi(y_1 \ldots y_{i-1} 0 y_i \ldots y_{n-1})}{\pi(y_1 \ldots y_{i-1} 1 y_i \ldots y_{n-1})} \leq \frac{1 + \delta}{1 - \delta}.$$
Call a distribution $\delta$-balanced if it is $\delta$-balanced at all positions $i = 1, \ldots, n$. Note that if the output distribution of a randomness source is $\delta$-balanced it is a Santha-Vazirani source (but not vice-versa).

Consider an arbitrary boolean function $f : \{0, 1\}^n \to \{0, 1\}$. Let $\pi_f^0$ be the probability that $f(x) = 0$ when $x$ is drawn according to the distribution $\pi$. That is, $\pi_f^0 = \sum_{x|f(x)=0} \pi(x)$. Similarly let $\pi_f^1 = \sum_{x|f(x)=1} \pi(x)$.

(b) Show that for every $f : \{0, 1\}^n \to \{0, 1\}$, and every $\delta \in [0, 1]$, there exists a $\delta$-balanced distribution $\pi$ over $\{0, 1\}^n$ such that $|\pi_f^0 - \pi_f^1| \geq \delta$.

(Hint: Consider separately the functions $f$ for which $|\mathcal{U}_f^0 - \mathcal{U}_f^1| \geq \delta$ and those for which $|\mathcal{U}_f^0 - \mathcal{U}_f^1| < \delta$, where $\mathcal{U}$ is the uniform distribution over $n$-bit strings.)

Conclude that there are no simple (deterministic) extractors which can extract a single $\epsilon$-balanced bit from all $\delta$-balanced Santha-Vazirani sources, with $\epsilon < \delta$.

**Problem 5:**

(a) (Randomized rounding.) Given a probability distribution $\rho$ over $R$ and random variable $X$, with range $[0, 1]$, define a probability distribution $\pi$ over $S = R \times \{0, 1\}$ as follows:

$$\pi((r, 1)) = \rho(r) \cdot X(r) \quad \text{and} \quad \pi((r, 0)) = \rho(r)\left[1 - X(r)\right]$$

Verify that $\pi$ is indeed a valid probability distribution. Now define a binary random variable $Z$ (i.e., with range $\{0, 1\}$), with underlying probability distribution $\pi$, as $Z(r, 0) = 0$ and $Z(r, 1) = 1$ for all $r \in R$. Show that $\mathbf{E}[Z] = \mathbf{E}[X]$.

(That is, instead of the real variable $X$, the binary random variable $Z$ can be used without changing the expectation (though the variance could increase). This is called randomized rounding because $Z$ can be considered to be sampled as follows: draw a sample from $X$, and using the value obtained as the bias, flip a coin, to get a rounded (0-1) value.)

(b) (Deterministic rounding.) Let $X$ be as above. Consider a new random variable $Z^*$ defined over $R$ and with respect to the same probability distribution $\rho$, as follows: $Z^*(r) = 1$ if $X(r) > \frac{1}{2}$ and 0 otherwise.

Using Markov’s inequality, show that $2\mathbf{E}[X] - 1 \leq \Pr[Z^* = 1] \leq 2\mathbf{E}[X]$. Conclude that if $\mathbf{E}[X] > 7/8$ then $\Pr[Z^* = 1] > 3/4$ and if $\mathbf{E}[X] < 1/8$ then $\Pr[Z^* = 1] < 1/4$.

(c) (Eliminating an auxiliary random source.) In this problem we consider a randomized algorithm $A$ which draws its randomness from two independent random sources, a “main” source (with an arbitrary distribution) and an auxiliary perfect random source. Our goal is to change it to an algorithm $B$ which uses only the main source, by enumerating over all random strings from the auxiliary source (while drawing only as many bits as $A$ draws from the main source).

Describe $B$ so that if the probability of error of $A$ is at most $1/8$ (when run using the two sources), then the probability of error of $B$ is at most $1/4$ (when run using only the main source). Prove that $B$ has these properties. (Hint: Use part (b). What should the real variable $X$ be?)

**Problem 6 (Extra Credit):**

In this problem we use basic linear algebra to analyze (weak) extraction from an SV source (see Lecture 15).

(a) (Collision probability.) Define a probability distribution $\pi$ over $\{0, 1\}^d$. We will view $\pi$ as a real vector of length $2^d$ (i.e. $\pi \in \mathbb{R}^{2^d}$), such that (with elements indexed by $i \in \{0, 1\}^d$) $\pi_i = \pi(i)$. Define collision probability of $\pi$, $\text{col}(\pi)$ to be the probability that two strings drawn independently according to $\pi$ are the same. Show that $\text{col}(\pi) = \|\pi\|^2$, where $\|v\|$ is defined as $\sqrt{\langle v, v \rangle}$.

(b) (An orthonormal basis.) Define $2^d$ vectors $\rho^{(s)}$ (for $s \in \{0, 1\}^d$) as follows: $\rho^{(s)} = \frac{1}{\sqrt{2^d}}(-1)^{\langle s, j \rangle}$. Note that $\|\rho^{(s)}\| = 1$, and each element in $\rho^{(s)}$ is $\pm \frac{1}{\sqrt{2^d}}$, the sign depending on whether $\langle s, j \rangle$ is even or odd. Show that $\langle \rho^{(s)}, \rho^{(t)} \rangle = 0$ for all $s \neq t$. 

(Hint: \( s \neq t \) means there is at least one position where the vectors \( s \) and \( t \) differ. Use this to show that all the vectors can be partitioned into pairs \((j_0, j_1)\) such that the parities of \(\langle s, j_0 \rangle \) and \(\langle t, j_0 \rangle \) are equal, and those of \(\langle s, j_1 \rangle \) and \(\langle t, j_1 \rangle \) are different.)

Hence these \(2^d\) vectors form an orthonormal basis for the vector space \(\mathbb{R}^{2^d}\). This basis is called the Fourier Basis.

(c) (Change of basis.) Recall that given an orthonormal basis any vector \( v \) can be written as a linear combination of the basis vectors, with the coefficients being the inner product of the vector \( v \) with basis vectors. So we can write \( \pi = \sum_s \langle \pi, \rho(s) \rangle \rho(s) \). Use this to rewrite \( \|\pi\|^2 \).

(d) Consider the extractor which on input \( r \in \{0,1\}^d \) and seed \( s \in \{0,1\}^d \) outputs the bit \( \langle r, s \rangle \). We consider feeding the extractor an input drawn according to the distribution \( \pi \). For each seed value \( s \), define \( \text{Gap}_s = \Pr_{r \sim \pi}[\langle r, s \rangle = 0] - \Pr_{r \sim \pi}[\langle r, s \rangle = 1] \). Show that \( \text{Gap}_s = \langle \pi, \rho(s) \rangle \).

(e) Deduce that \( \mathbb{E}_{u \sim \mathcal{U}_{\text{d}}}[\text{Gap}_s^2] = \text{col}(\pi) \), where \( \mathcal{U}_{\text{d}} \) is the uniform distribution over \(\{0,1\}^d\).

(f) From this, using the fact that \( \mathbb{E}[X]^2 \leq \mathbb{E}[X^2] \), conclude that

\[
\left| \Pr_{r \sim \pi, s \sim \mathcal{U}_{\text{d}}}[\langle r, s \rangle = 0] - \Pr_{r \sim \pi, s \sim \mathcal{U}_{\text{d}}}[\langle r, s \rangle = 1] \right| \leq \|\pi\|.
\]

Note that the left hand side is the bias of the extracted bit, when the input \( r \) is drawn according to the distribution \( \pi \) and the seed \( s \) is drawn independently from \( \mathcal{U}_{\text{d}} \). Finally, show that when \( \pi \) is an SV source with bias bounded by a constant less than 1, \( \|\pi\| = 2^{-O(d)} \).