1. The covariance between two random variables X and Y is given by

$$\mathrm{Cov}(X, Y) = E\left[(X - E(X))(Y - E(Y))\right].$$

The random variables are said to be *uncorrelated* if $\mathrm{Cov}(X, Y) = 0$. "If X and Y are independent, then they should be uncorrelated." Is this statement true? If true, prove the statement. If not, provide a counterexample where the statement does not hold. **(2 points)**

2. Consider two events E_1 and E_2. Let X, Y denote indicator random variables corresponding to the events E_1 and E_2, respectively. i.e.,

$$X = \begin{cases} 1 & \text{if } E_1 \text{ occurs} \\ 0 & \text{otherwise} \end{cases}, \quad Y = \begin{cases} 1 & \text{if } E_2 \text{ occurs} \\ 0 & \text{otherwise} \end{cases}.$$

Show that the random variables X and Y are independent if and only if the events E_1 and E_2 are independent. **(2 points)**

3. Roll a fair six-sided die. Suppose X denotes the number appearing on the top of the die. Compute the mean and variance of X. **(3 points)**

4. Imagine that you are a researcher studying a new virus. In a village of 100 people, you know 10 of them are infected, but you do not know who they are. To study the virus, you need at least one infected person's blood. Each test costs \$10. You want to spend as little money as possible while being 90% sure you will find at least one infected person. How many people should you test? (*Hint*: To simplify calculations, assume you pick each person uniformly at random, with replacement. Although sampling without replacement is more realistic, it is harder to calculate.)

(3 points)

5. Let X be a continuous uniform random variable on the interval $[0, 1]$. You are interested in obtaining an exponential random variable Y with mean 0.5 by viewing Y as a function of X. How will you do this? (*Hint*: Use CDFs) **(3 points)**

6. Let $A \in \mathbb{F}^{m \times n}$. Let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$ denote the singular values of A. Show that:

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq \sigma_1, \qquad \forall \mathbf{x} \neq \mathbf{0}.$$

Suggest an $\mathbf{x}^*$ for which:

$$\frac{\|A\mathbf{x}^*\|}{\|\mathbf{x}^*\|} = \sigma_1.$$

(3 points)

7. You are given a coin with unknown bias θ, i.e., probability of getting Heads in a coin toss is θ. You perform n independent coin tosses and have associated random variables X_i, where:

$$X_i = \begin{cases} 1 & \text{toss-}i \text{ results in Heads} \\ 0 & \text{otherwise} \end{cases}.$$

Consider the estimator $\hat{\Theta}_n \triangleq \frac{X_1 + X_2 + \cdots + X_n}{n}$ for the unknown θ. How many tosses n are sufficient so that: $P(|\hat{\Theta}_n - \theta| \geq 0.01) \leq 0.02$? **(3 points)**

Hint: You may use Chebyshev's inequality or Hoeffding's inequality here.

8. This question has three parts. There is a sender who sends a signal $X = +1$ or $X = -1$, with probabilities 0.9 and 0.1, respectively. The receiver receives $Y = X + Z$, where $Z \sim N(0, 0.25)$. Suppose the receiver obtained $Y = y$.

   (i) Obtain the optimal rule/test that computes $\hat{X}$ using y such that $P(\hat{X} \neq X \mid Y = y)$ is minimized. i.e., obtain the constant τ such that:

   $$\hat{X} = \begin{cases} +1 & \text{if } y \geq \tau, \\ -1 & \text{otherwise.} \end{cases}$$

   (3 points)

   (ii) For the derived optimal rule in part (i), compute the error probabilities:

   $$P(\hat{X} \neq X \mid X = +1) \quad \text{and} \quad P(\hat{X} \neq X \mid X = -1).$$

   Hint: Normalize the random variables appropriately to make them standard normal and use the data in Table 1. **(2 points)**

z	-2.549	-1.764	-1.451	0	1.451	1.764	2.549
$\Phi(z)$	0.0054	0.0389	0.0734	0.5	0.9266	0.9611	0.9946

Table 1: Table of CDF of the standard normal random variable

   (iii) Compute the error probability $P(\hat{X} \neq X)$. **(1 points)**