Assignment 1

Rohan (241110057)

Question 1

It is promised that a given coin is either fair, i.e., $P(\text{Head}) = \frac{1}{2}$, or biased, i.e., $P(\text{Head}) = \frac{1}{2} + \epsilon$, where $0 < \epsilon < \frac{1}{2}$.

We are tasked with showing that using $\frac{100}{\epsilon^2}$ coin tosses, it is possible to correctly determine the type of coin (whether it is fair or biased) with at least $\frac{4}{5}$ probability. Specifically, we need to develop an algorithm that will require at most $\frac{100}{\epsilon^2}$ tosses and will have the following guarantees:

- If the coin is fair, the algorithm will return "fair" with probability at least $\frac{4}{5}$.
- If the coin is biased, the algorithm will return "biased" with probability at least $\frac{4}{5}$.

Solution

Consider a coin that can either be:

- **Fair**: $P(\text{Head}) = \frac{1}{2}$
- **Biased**: $P(\text{Head}) = \frac{1}{2} + \epsilon$, where $0 < \epsilon < \frac{1}{2}$

Number of Tosses

Let $n = \frac{100}{\epsilon^2}$ denote the number of tosses. Define X as the number of heads observed in these n tosses.

Expected Values

For a fair coin, the expected number of heads is:

$$\mathbb{E}[X] = n \times \frac{1}{2} = \frac{n}{2}$$

For a biased coin, the expected number of heads is:

$$\mathbb{E}[X] = n \times \left(\frac{1}{2} + \epsilon\right) = \frac{n}{2} + n\epsilon$$

Variance and Distribution

Each coin toss is independent, and the number of heads X follows a binomial distribution. Using the Central Limit Theorem (CLT), we can approximate the distribution of X for large n as a normal distribution.

- If the coin is fair:

$$X \sim \mathcal{N}\left(\frac{n}{2}, \frac{n}{4}\right)$$

- If the coin is biased:

$$X \sim \mathcal{N}\left(\frac{n}{2} + n\epsilon, \frac{n}{4}\right)$$

Hypothesis Testing

We aim to design a hypothesis test to determine whether the coin is fair or biased based on the observed number of heads X. Our goal is to distinguish between the two cases with high probability, ensuring that the probability of error is at most $\frac{1}{5}$.

Hoeffding's Inequality

Hoeffding's inequality provides a bound on the probability of large deviations from the expected value. For the binomial distribution X:

$$P\left(|X - \mathbb{E}[X]| \geq t\right) \leq 2\exp\left(-\frac{2t^2}{n}\right)$$

We seek to choose t such that the probability of error is at most $\frac{1}{5}$.

Threshold Calculation

To distinguish between the fair and biased coins, we set a threshold $t = n\epsilon$, which ensures that we can separate the two hypotheses with high probability. Using Hoeffding's inequality, we can bound the probability that X deviates from its expected value by more than $n\epsilon$:

$$P\left(|X - \mathbb{E}[X]| \geq n\epsilon\right) \leq 2\exp\left(-2n\epsilon^2\right)$$

To guarantee a confidence level of at least $\frac{4}{5}$, we require:

$$2\exp\left(-2n\epsilon^2\right) \leq \frac{1}{5}$$

Taking log on both sides:

$$-2n\epsilon^2 \leq \ln\left(\frac{1}{5}\right) = -\ln(5)$$

Thus, the number of tosses must satisfy:

$$n \geq \frac{\ln(5)}{2\epsilon^2}$$

Since $\ln(5) \approx 1.609$, we have:

$$n \geq \frac{1.609}{2\epsilon^2} \approx \frac{100}{\epsilon^2}$$

This shows that $n = \frac{100}{\epsilon^2}$ tosses are sufficient to achieve the desired probability of success.

Algorithm

- Toss the coin $n = \frac{100}{\epsilon^2}$ times.
- Let X be the number of heads observed.
- If $X \leq \frac{n}{2} + n\epsilon$, we conclude that the coin is fair.
- If $X > \frac{n}{2} + n\epsilon$, we conclude that the coin is biased.

This algorithm makes sure that with at most $\frac{100}{\epsilon^2}$ tosses, we can determine whether the coin is fair or biased with at least $\frac{4}{5}$ probability.