1. In the (Uncapacitated) Metric Facility Location problem, we are given as input a set \( D \) of clients and a set \( F \) of facilities. Each facility \( i \in F \) has an opening cost \( f_i \). For any \( i, j \in (F \cup D) \) there is a distance \( c_{ij} \) between them. We assume the distances form a metric; that is, \( c_{ii} = 0 \), \( c_{ij} = c_{ji} \), and \( c_{ij} + c_{jk} \geq c_{ik} \) for all \( i, j, k \in (F \cup D) \). The goal is to find a subset \( S \subseteq F \) of facilities to open such that the following cost is minimized:

\[
\text{cost}(S) = \sum_{i \in S} f_i + \sum_{j \in D} c(j, S),
\]

where \( c(j, S) = \min_{i \in S} c_{ij} \) is the connection cost of client \( j \in D \) and \( S(j) := \arg \min_{i \in S} c_{ij} \) is the facility that client \( j \) is connects to.

In the class, we learned a 6-approximation using linear programming and rounding (See Anupam Gupta’s (CMU) lecture notes): We first obtained an optimal LP solution \((x^*, y^*)\); and converted \((x^*, y^*)\) into another feasible solution \((x', y')\). This was done by defining \( \Delta_j := \sum_{i \in V} c_{ij}x_{ij}^* \). Then for each client \( j \in D \) and a facility \( i \in F \) we set \( x'_{ij} = 2x_{ij}^* \) if \( c_{ij} \leq 2\Delta_j \) and otherwise \( x'_{ij} = 0 \). We then set \( y'_i = \max_j x'_{ij} \). We argued that \((x', y')\) is a feasible LP solution and it has cost at most twice that of the optimal LP solution.

Then, we defined the neighborhood \( N_j \) of a client \( j \) to be all facilities within distance \( 2\Delta_j \) of \( j \). The algorithm then performed the following steps:

- Find the client \( j \) such that \( \Delta_j \) is minimized
- Open the cheapest facility \( i \) in \( N_j \)
- Let \( E_j \) be the extended neighborhood of \( j \) which in includes any client \( j' \) where \( N_j \cap N_{j'} \neq \emptyset \)
- Assign all clients in \( E_j \) to \( i \)
- Remove all clients in \( E_j \) and all facilities from \( N_j \)
- Recurse

Intuitively, one can think of the facilities in \( N_j \) as the facilities contained in a ‘ball’ of radius \( 2\Delta_j \) around client \( j \). Then, the clients in \( E_j \) are clients we remove because their ‘balls’ (of radius \( 2\Delta_{j'} \) for \( j' \in E_j \)) intersect \( j \)’s balls.

Now one may one wonder why we remove any client \( j' \) whose ball intersects \( j \)’s ball and, instead, didn’t just remove clients that are ‘inside’ \( j \)’s ball (that is, clients \( j' \) where \( c_{jj'} \leq 2\Delta_j \)). Consider altering the algorithm in the following way that formalizes this intuition.

- Find the client \( j \) such that \( \Delta_j \) is minimized
• Open the cheapest facility \( i \) in \( N_j \)
• Let \( C_j \) be the close neighborhood of \( j \) which includes any client \( j' \) where \( c_{jj'} \leq 2\Delta_j \)
• Assign all clients in \( C_j \) to \( i \)
• Remove all clients in \( C_j \) and all facilities from \( N_j \)
• Recurse

This algorithm does not give an \( O(1) \)-approximation. Explain why.

2. Assume that you have a fair coin that yields either a head or tail, each with probability 1/2. Let \( n > 0 \) be a parameter. For simplicity, assume that \( \log n \) and \( n \) are both integers. Repeat flipping the coin \( k \) times sequentially. What is the (asymptotically) minimum \( k \) where you observe at least one head with probability at least \( 1 - \frac{1}{n} \)? Also, what is the minimum \( k \) where you observe at least \( \log n \) heads with probability at least \( 1 - \frac{1}{n} \)?