1. (5p.) Let us consider the maximum cut problem. Given a graph $G = (V, E)$ with n nodes and m edges, we want to partition the nodes in two subsets A and B such that the number of edges in $A \times B$ (ie. the number of edges with an endpoint in each set) is maximized. The partition $V = A \cup B$ is called a *cut* of G and the number of edges with endpoints in each set is referred to as the number of edges in the cut.

   We use the following randomised algorithm:

   - Each node $u \in V$ picks a value in $\{0, 1\}$ uniformly at random,
   - If u picked 0 it joins set A, otherwise it joins set B.

   We want to analyse how good a cut this simple algorithm provides us.

   (a) (1p.) Prove that the probability that an edge i is in the cut is $\frac{1}{2}$.

Let's call the cut set as C and the maximum cut as $C_{\max}$. A cut edge is an edge whose endpoints belong to different sets A and B, while an edge is not a cut edge if its endpoints are both in A, or both in B. Let $edge_i = \{u, v\}$. Since each node is uniformly assigned at random 0 or 1, it means that

$\Pr[u \in A] = \Pr[u \in B] = \Pr[v \in A] = \Pr[v \in B] = \frac{1}{2}$. From this identity, we can calculate:

$\Pr[u \in A, v \in A] = \Pr[u \in A]\Pr[v \in A] = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. Similar for other identities

$\Rightarrow$ $\Pr[u \in A, v \in A] = \Pr[u \in A, v \in B] = \Pr[u \in B, v \in A] = \Pr[u \in B, v \in B] = \frac{1}{4}$

By definition, we have: $\Pr[edge_i \in C] = \Pr[u \in A, v \in B] + \Pr[u \in B, v \in A] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

$\Rightarrow$ $\Pr[edge_i \in C] = \frac{1}{2}$ (proven)

(b) (1p.) Let X_i be the random variables defined for each edge $i \in [m]$ s.t. $X_i = 1$ if the edge i has one endpoint in A and one endpoint in B, otherwise $X_i = 0$. Let X be the random variable giving the number of edges in the cut $A \cup B$. Prove that the expected number of edges in the cut is at least $m/2$, ie $\mathrm{E}[X] \geq \frac{m}{2}$.
hint: write X using X_i, $\forall i \in [m]$.

By the definition, we can find the expectation of X_i as follows:

$$E[X_i] = \frac{1}{2}\left(E[X_i \in C] + E[X_i \notin C]\right) = \frac{1}{2}(0+1) = \frac{1}{2}$$ (since each case occurs equally likely)

On the other hand, we have $X = X_1 + \ldots + X_m \forall i \in m$

⇨ $E[X] = \sum_{i=1}^{m} E[X_i] = \frac{m}{2}$ (linearity of expectation)

Now, let's define the probability of success as follows:

$p = \Pr\left[X \geq \frac{m}{2}\right]$. From this,

$$E[X] = \frac{m}{2} = \sum_{x<m/2} x \Pr[X = x] + \sum_{x \geq m/2} x \Pr[X = x]$$

⇨ $\frac{m}{2} \leq (1-p)\left(\frac{m}{2} - 1\right) + pm$

⇨ $p \geq \frac{1}{m/2+1}$. Because the probability of success is higher than failure, it follows that:

⇨ $E[X] \geq \frac{m}{2}$ (proven)

(c) (1p.) Show that the expected output of this algorithm is a 2-approximation for the maximum cut.

From part (b), we observe that $E\left[X_i\right] = \Pr\left[X_i \in C\right] = \frac{1}{2}(0+1) = \frac{1}{2}$

⇨ $E\left[\left|C\right|\right] = \sum_{i=1}^{m} E\left[X_i\right] = \sum_{i=1}^{m} \Pr\left[X \in C\right] = \frac{1}{2}m$ (from linearity of expectation)

And since $m \geq \left|C_{\max}\right|$ => $E\left[\left|C\right|\right] = \frac{1}{2}m \geq \frac{1}{2}\left|C_{\max}\right|$ (proven)

Therefore, the expected output of the algorithm isa 2-approximation for maximum cut

(d) (1p.) Use Markov's inequality to show that $\mathrm{P}\left(X \leq m\left(\frac{1}{2}-\varepsilon\right)\right) \leq 1-\varepsilon$, where $0 \leq \varepsilon < \frac{1}{2}$.

Hint: Use the indicator random variable Y_i for the event "edge i is not in the cut", and Y the number of edges not in the cut.

The Markov's inquality is $\Pr[X \geq \alpha] \leq \frac{E[X]}{\alpha}$, where α is some constant factor. Let's call Y a random variable of the number of edges not in the cut. => $Y = m - X$

- ⇨ $\Pr[Y \geq \alpha] \leq \frac{E[Y]}{\alpha}$ => $\Pr[(m-X) \geq \alpha] \leq \frac{E[m-X]}{\alpha}$
- ⇨ $\Pr[(m-X) \geq \alpha] \leq \frac{E[m-X]}{\alpha}$ => $\Pr[X \leq m-\alpha] \leq \frac{E[m-X]}{\alpha}$

Let $\alpha = m-(1/2-\varepsilon)m$, the inequality becomes:

- ⇨ $\Pr\left[X \leq \left(\frac{1}{2}-\varepsilon\right)m\right] \leq \frac{E[m-X]}{m-(1/2-\varepsilon)m} = \frac{m-E[X]}{(1/2+\varepsilon)m}$ (due to linearity of expectation)

From previous parts, we have $E[X] = \frac{m}{2}$

- ⇨ $\Pr\left[X \leq \left(\frac{1}{2}-\varepsilon\right)m\right] \leq \frac{m-m/2}{(1/2+\varepsilon)m} = \frac{m/2}{(1/2+\varepsilon)m} = \frac{1}{1+2\varepsilon} \leq 1-\varepsilon$, for $0 \leq \varepsilon < \frac{1}{2}$ (proven)