1. Lecture 2: let there be two hypotheses regarding the behavior of a coin:
   - \( H_0 \): upon a toss, the coin falls on “head” with probability \( p \)
   - \( H_1 \): upon a toss, the coin falls on “head” with probability \( q \) (\( q < p \))

   The coin was tossed \( N \) times, and fell on “head” \( h \) of those times.
   a. Write the likelihood ratio \( \Lambda(h) \) of the \( N \)-toss event described above.
   b. Prove that the decision rule comparing the likelihood ratio \( \Lambda(h) \) to a bound \( c \) is equivalent to comparing the number of heads \( h \) to a bound \( c^* \). In other words, there exists a number \( c^* \) such that accepting \( H_0 \) whenever \( h > c^* \) is equivalent to accepting \( H_0 \) whenever \( \Lambda(h) > c \).

   Let \( N = 5 \), \( p = 0.75 \), \( q = 0.33 \). What value of \( c^* \) should be used if:
   c. The false negative probability should be close to but not exceed 0.016?
   d. The false positive probability should be close to but not exceed 0.05?

2. Lecture 3: let \( T \) be a full, directed, binary tree of depth \( k > 1 \). Denote by \( G \) the graph obtained when adding directed edges from all the leaves of \( T \) to its root.
   a. Prove that nodes of equal distance from the root have equal PageRank in \( G \) (although \( G \) is not a tree, we still refer to the root of \( T \) as “the root”).
   b. Denote by \( p_j \) the PageRank (in \( G \)) of nodes at distance \( j \) from the root. With this notation, \( p_0 \) denotes the root’s PageRank. Write an expression for \( p_j - p_{j+1} \) as a function of \( p_0 \), \( N \) (the total number of nodes), and \( d \) (PageRank’s probability of following an outlink).

3. Lecture 3: let \( H = (V,E) \) be a directed Web graph, and let there be two disjoint page sets \( G \) (for “good”) and \( B \) (for “bad”) such that \( V = G \cup B \). The pages in \( B \) might be pages of spammers, and we would like to take advantage of knowing their identities in order to better rank the good pages of the set \( G \).

   Let \( W \) denote the (binary) adjacency matrix of \( H \), and let \( M \) denote the adjacency matrix of \( H \) where the entry corresponding to any link between pages of \( G \) and \( B \) (in either direction) is -1 instead of 1. Note that there is no change to the entries corresponding to links within \( G \) or within \( B \) – they remain 1.

   Prove that computing HITS authority scores using the matrix \( M \) will result in all \( G \)-pages receiving the same authority scores as when computing HITS authority scores using \( W \). In other words, this method is of no help since it does not affect the scores of good pages with bad neighbors. Hint: prove a transformation between the principal eigenvectors of \( M^TM \) and \( W^TW \).
4. Lecture 4: let \( G=(V,E) \) be a weighted directed irreducible graph. For node \( v \in V \), \( \text{let} \quad d_{\text{in}}(v) \quad \text{denote the weighted in-degree of} \ v \quad \text{in} \ G \). Let \( W \) denote the sum of all link weights in \( G \), and let \( N=|V| \).

For some \( \beta \in (0,1) \), let us change the definition of SALSA’s random walk on the authority chain of \( G \) as follows: upon leaving node \( x \), SALSA will perform - with probability \( \beta \) - a normal co-citation step, and with probability \( 1-\beta \) will jump to some node \( v \in V \), chosen at random proportionally to \( d_{\text{in}}(v) \). Write (and prove) a closed form expression for node \( x \)'s resulting authority score, using \( \beta, d_{\text{in}}(x) \) and \( W \).

Note: by “co-citation step” we refer to the normal SALSA step of retreating from \( x \) to a node \( k \) that links to \( x \) (i.e. backing up on an inlink of \( x \)), and then leaving \( k \) through an outlink.