Problem 1 (Adding a Link to a Network)**.** Fix an undirected network. For each of the following network statistics, when a link is added to the network, does the statistic always increase, always decrease, or sometimes increase and sometime decrease (depending on the network and the location of the new link)? For each answer, either give a proof that the statistic always increases or always decreases, or an example that shows it can go either way.

(a) average degree

(b) diameter

(c) average path length

(d) overall clustering coefficient

(e) average clustering coefficient

(f) decay centrality, for an arbitrary node i

(g) betweenness centrality, for an arbitrary node i

Problem 2 (The Adjacency Matrix)**.** Let $\mathbf{g}$ be the adjacency matrix of an undirected network, and let $\mathbf{1}$ be the column vector whose elements are all 1. In terms of these quantities write expressions for:

(a) The vector $\mathbf{d}$ whose elements are the degrees d_i of the nodes.

(b) The number m of edges in the network.

(c) The matrix $\mathbf{N}$ whose element N_{ij} is the number of common neighbors of nodes i and j.

(d) The total number of triangles in the network, where a triangle means three nodes, each connected by edges to both of the others.

Problem 3 (Betweenness Centrality in Trees)**.** Consider an undirected (connected) tree of n vertices. Suppose that a particular vertex k in the tree has degree d, so that its removal would divide the tree into d disjoint regions, and suppose that the sizes of those regions are $n_1, \ldots, n_d$.

(a) Show that the betweenness centrality of the vertex is

$$B_k = 1 - \sum_{m=1}^{d} \frac{n_m(n_m - 1)}{(n-1)(n-2)}.$$

(b) Using this result, calculate the betweenness of the ith vertex from the end of a "line graph" of n vertices, i.e., n vertices in a row.

Problem 4 (Expected Degree)**.** First, some definitions. Fix an undirected graph $G = (N, E)$ with finitely many nodes, none of which have degree zero. We will use $N(i)$ to denote the *neighborhood* of a node i—that is, the set of all nodes j that share some edge with i. Define d_i to be the degree of node i, which is also the size of i's neighborhood: $d_i = |N(i)|$.[1] Let $P(d)$ be the fraction of the nodes in the graph with degree d.

(a) Suppose we pick an edge uniformly at random[2] and then pick either node of that edge with equal probability. Call that node i. Let D be the degree of i; it is a random variable because the node was random. What is the expectation of D? Write your answer in terms of P.

(b) Prove that the expectation of D is at least as large as the mean of P.

(c) We make a definition to keep track of how popular i's neighbors are, on average.

Definition. Define δ_i to be the arithmetic mean of the degree of i's neighbors. That is,

$$\delta_i = \frac{\sum_{j \in N(i)} d_j}{d_i}.$$

Theorem. For graph $G = (N, E)$ satisfying the conditions given in this problem,

$$\frac{1}{|N|} \sum_{i \in N} d_i \leq \frac{1}{|N|} \sum_{i \in N} \delta_i.$$

Prove this statement.

[1] Recall we use the notation $|S|$ to denote the number of elements in the set S.
[2] That is, each edge of the graph is chosen with equal probability.

14.15 / 6.207 Networks
Spring 2022