Math 420, Spring 2020
Fourth Team Homework

Exercise 1. Consider an Erdös-Rényi random graph G with $n = 100$ vertices and probability $p = 0.9$ for each edge, i.e. $G \in \mathcal{G}_{100,0.9}$.

1. (1pt) What is the expected number of edges ?
2. (1pt) For each vertex v, the degree $deg(v)$ is defined as the number of edges that have v as one of end points. (Thus in a complete graph with n verticies, each vertex has degree $n-1$). For the random graph G, compute the expected degree of each vertex.
3. (1pt) Assume each edge of G is colored either in red, or in blue. Given an edge, assume the probability of being red is 30% whereas the probability of being blue is 70%. Determine the expected numbers of red edges and of blue edges.
4. (1pt) Detemine the expected number of 3-cliques.
5. (1pt) Detemine the expected number of 4-cliques.

Exercise 2. (5pts) Write a Matlab script (call it teamXhw4.m) and a Matlab function (call it cliques.m) that perform the following steps:

1. [2pts] The Matlab function cliques.m should have the following preamble:
function qp = cliques(E,n,m,p)

```
% CLIQUES computes cumulative count of cliques of a given ordered graph.
% Use:
%  qp = cliques(E,n,m,3)
% Inputs:
%  E is a mx2 matrix of ordered edges (matrix of integers).
%  n  is the number of vertices (integer)
%  m is the number of edges (integer)
%  p is the clique order (integer); for now p should be set to 3.
% Outputs:
%  qp is a m-vector of integers. qp(k) should contain the number of p-cliques
%       of the graph whose set of edges is E(1:k,1:2)
%
```

The function should compute the cumulative count of 3-cliques of a given graph. Its inputs are:

1. the ordered sequence of unoriented edges $E(1:m, 1:2)$, where the k^{th} edge is between vertex $E(k,1)$ and $E(k,2)$;
2. the number of vertices n;

3. the number of edges m (somewhat redundant, since it can be read from $size(E, 1)$);

4. the clique order p, that should be set to 3 for now. (In the subsequent homework you will be asked to implement the case $p = 4$).

Its output is a length m vector of integers, call it qp, that has the following significance: $qp(k)$ is the number of p-cliques (3-cliques in this homework) in the graph over n vertices that has k edges, namely $E(1 : k, 1 : 2)$, that is, $\{(E(1,1), E(1,2)), (E(2,1), E(2,2)), (E(3,1), E(3,2)), \ldots, (E(k,1), E(k,2))\}$.

In this homework the function cliques() should work for $p = 3$.

2. The Matlab script teamXhw4.m should do the following:

1pt Load data file and turns into an ordered sequence of edges. If the raw data is a weight matrix (such as BKFRAT), then use the weights to order the edges (from highest to smallest weight); if data set is sequence of links (such as in the ca-*** arXiv preprint list server), then use that sequence of links as the ordered sequence of edges. For testing purposes use BKFRAT until you get a different dataset to work with. See description of BKFRAT in the Discovery Lecture 1, " From Data to Weighted Graphs and Graph Laplacian".

1pt Call cliques() Matlab function you wrote at part 1) to construct the sequence of 3-cliques on this ordered edges.

1pt Plot the cumulative count of 3-cliques $qp(1 : m)$ against the number of edges $1 : m$.