MAT331 Project 3 – The Giant Component of a Random Graph

In this project you will explore "giant component" of a random graph. Recall that the vertices of an graph G can be partitioned into sets called connected components, such that every pair of vertices in each set can be joined by a path consisting of edges of the graph, and no pair of vertices from different sets can be joined by such a path. The number, size, and structure of the connected components are important properties of graphs that arise when modeling networks. You will investigate the relationship between the connected components (particularly the size of the largest one - the giant component) in a random graph and the number of edges in the graph. With few edges, all the connected components are small and there are many of them. As the number of edges increases, eventually it is very likely that the graph will be connected (i.e. have a single connected component). But between these two extremes there is a transition point when the graph starts to have one large component, and many smaller ones. You will empirically explore this transition point.

1. Write a function `randomGraph(n, p)` that generates the adjacency list of a random graph with n vertices, where each edge occurs with probability p.

2. Plot visual representations of the random graphs for $n = 10, 20, 30, 40$ and 50. (Look at hints given in the lectures).

3. Write a function `largestComponent(graph)` which returns the size of the largest component of the graph, where the input is in adjacency list form. Demonstrate that you understand what the code does, by working out in detail (line by line) what happens when you call `largestComponent([[1],[0,2],[1,3,4],[2],[2]])`.

4. Write a function `expectedLargest(n,p, trials)` which generates `trials` random graphs and finds the average normalized size of the largest component. The normalized size of a component is the number of vertices in the component divided by n, the total number of vertices of the graph.

5. For $n = 20$, what is the expected normalized size of a random graph with $p = t/n$, where $t = 1$? Do the same for the range $t = 0$ to $t = 4$ in increments of 0.1. Then make a plot of the expected normalized size (y axis) vs t (x axis).

   Finally, do the same thing for $n = 20, 60, 120, 300, 600, 1000$ and put all plots on the same axes, one curve for each n.

6. Interpret the results about the size of the giant component you computed in (5). What happens at $t = 1$? What do you think would happen if you were able to test random graphs with very large n?

You will submit the following on brighspace: Include the following:

- A .ipynb notebook file named `code_firstname_lastname.ipynb` with all your code from the Programming tasks section. Make sure to clearly separate the various parts. Also, test this notebook by restarting it and running from the beginning; this should produce no errors.

- A pdf file named `report_firstname_lastname.pdf` with your work for questions 6. The exact format of this report is up to you. It can be hand-written (legibly) and then scanned (legibly) into a pdf document. Or it can be be made on a computer. It must be in pdf form (it is generally easy to convert other forms such as .doc to pdf).