Project 6

This project focuses on random graphs and configuration models.

1) Analysis of Random Graphs $G(n,p)$

Fix an integer $n \geq 10,000$. For various values of the edge probability p such that pn is a constant, use computational experiments (plots and calculations) to verify the following theoretical properties of $G(n,p)$:

a) Mean degree c

   – Verify that the mean degree $c = p(n-1)$.

   – Plot c as a function of p.

b) Degree Distribution p_k: show that p_k follows a Poisson distribution $p_k = e^{-c}\frac{c^k}{k!}$ by plotting the empirical degree distribution (histogram) and overlay the theoretical Poisson curve.

c) Clustering Coefficients: verify that both the local and global clustering coefficients of $G(n,p)$ are equal to p.

d) Giant Component Threshold: confirm that the threshold probability for the emergence of a giant component is $\frac{1}{n-1}$.

e) Fraction S of Vertices in the Giant Component: show that the fraction S of vertices in the giant component satisfies: $1 - S = e^{-cS}$ by comparing the empirical value of S with the theoretical prediction (using numerical methods if needed).

f) Small Components:

– Verify that small components are trees.

– Show that the average size of small components is:

$$R = \frac{2}{2 - c + cS}$$

g) Fraction of Vertices in Small Components: verify that the fraction of vertices in small components follows

$$\frac{e^{-sc}(sc)^{s-1}}{s!}$$

h) Diameter: show that the diameter of $G(n,p)$ follows:

$$\text{diameter} = A + \frac{\ln n}{\ln c}$$

where A is a constant.

2) Empirical Analysis of Real-World Network

Analyze the `ca-GrQc` dataset from https://snap.stanford.edu/data/ca-GrQc.html in the Stanford Large Network Dataset Collection.

a) Construct a graph G based on the data set and analyze some basic network properties of G including order, size, density, connectivity (if G is not connected, find the number of components of G and the fraction of vertices in the largest component), and clustering coefficient.

b) Generate a configuration model G^* that has the same degree sequence as that of G's.

c) Analyze some basic network properties of G^*.

d) Identify similarities and differences between G and G^*.