

Selected Topics in Frontiers of Statistics Third Assignment

12111620 Yixuan Ding

May 4, 2024

Part 1

Q1

For question 1, networkx is utilized to generate ER network. As the question requests, the average degree of the network is 0.5. The network size is chosen to be 10000, thus the probability for connection is set to 0.00005.

After generating graph, the component of the graph can be obtained, so as their sizes. Calculated the number for different sizes and gain the distribution of component size. The graph is shown below:

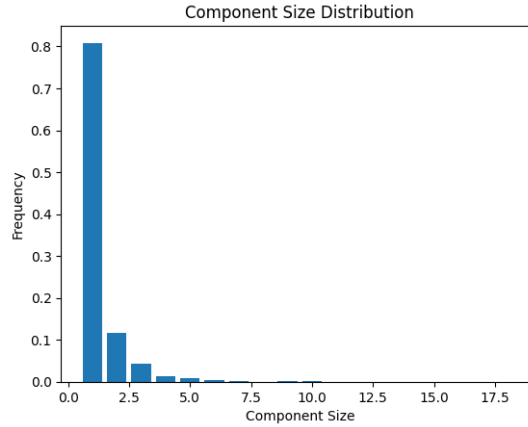


Figure 1: Simulated Component Size Distribution

Q2

To calculate P_s numerically, several formulation needs to be introduced:

- $G_0(x) = \sum_{k=0}^{\infty} p_k x^k$: Generating function for degree distribution.
- $G_1(x) = \frac{G'_0(x)}{G'_0(1)}$: Generating function for second layer neighbor degree distribution.
- $H_1(x) = xG_1(H_1(x))$: Generating function for size of component reached by choosing a random edge and following it to one of its end.
- $H_0(x) = xG_0(H_1(x))$: Generating function for size of component located by randomly choosing a vertex.

For p_k of G_0 , it can be represented by the frequency of degree in graph generated above. Then G_1 can be obtained by G_0 . For the calculation of H_1 , it is noticed that the equation is of self-consistency

form, so that iteration can be utilized by setting H_1 to x at the beginning. After calculating H_1 , H_0 can be derived.

Finally, P_s can be derived by

$$P_s = \frac{1}{s!} \left. \frac{d^s H_0}{dz^s} \right|_{z=0}$$

The distribution of P_s is shown below:

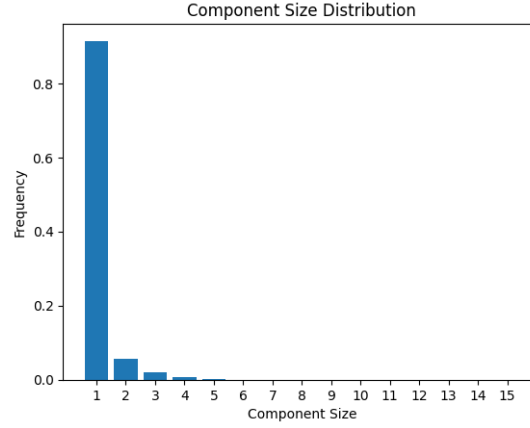


Figure 2: Calculated Component Size Distribution

Comparison

The simulated component size distribution is nearly the same as the theoretic calculation of the distribution. They all have more amount of component when the size is smaller, and the number of component decays as the size growing. However, theoretically, there are more small components than simulated scenario.

Part2

Q3

For poisson degree distribution with average degree $\langle k \rangle$, S and T can be calculated separately and expressed by $\langle k \rangle$.

$$\begin{cases} S = 1 - e^{-\frac{S}{\langle k \rangle}} \\ T = \frac{1}{\langle k \rangle} \end{cases}$$

Then $S = 1 - e^{-\frac{S}{\langle k \rangle}}$, and the plot is derived:

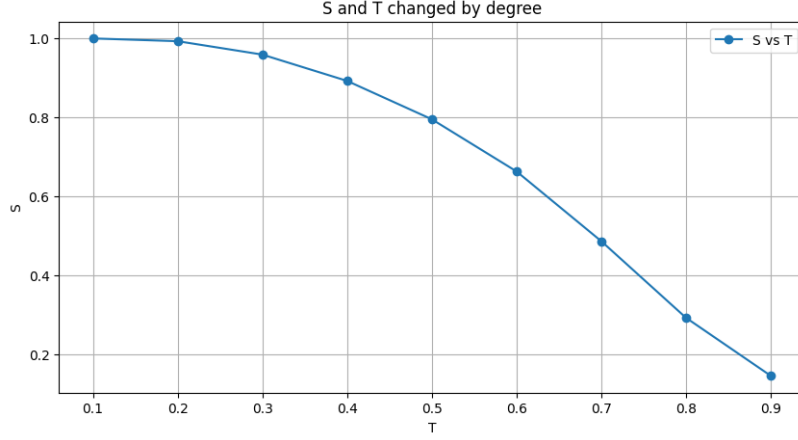


Figure 3: S vs T

Q4

For power law degree distribution and $p(k) \propto k^{-2.5}$, the minimum degree and the maximum degree of the network is 2 and \sqrt{N} . Here N is set to 1001. First, G_0 is calculated by degree distribution. After that we can obtain G_1 , H_1 and H_0 as before. Notice that:

$$\begin{cases} S = 1 - G_0(u) \text{ and } u \equiv H_1(1) \\ T = \frac{G'_0(1)}{G''_0(1)} \end{cases}$$

By calculation:

$$\begin{cases} T = 0.17045 \\ S = 0.65918 \end{cases}$$

Appendix

More relevant detailed information you can find in: **Github Repository**