Homework 2

Name: Yang Sun, Andrew ID: yksun

Q2:

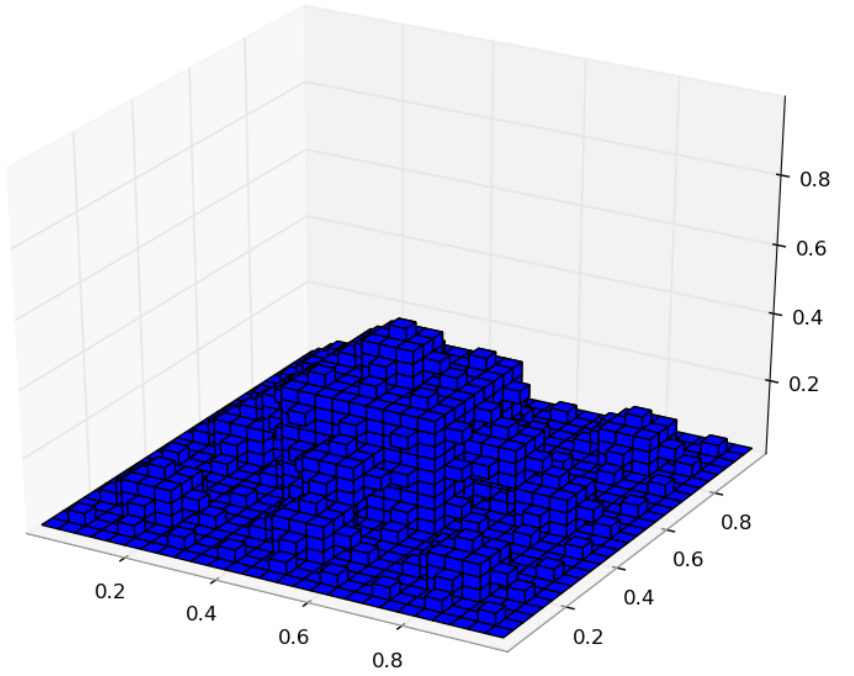

Figure 1: image of the fractal K3

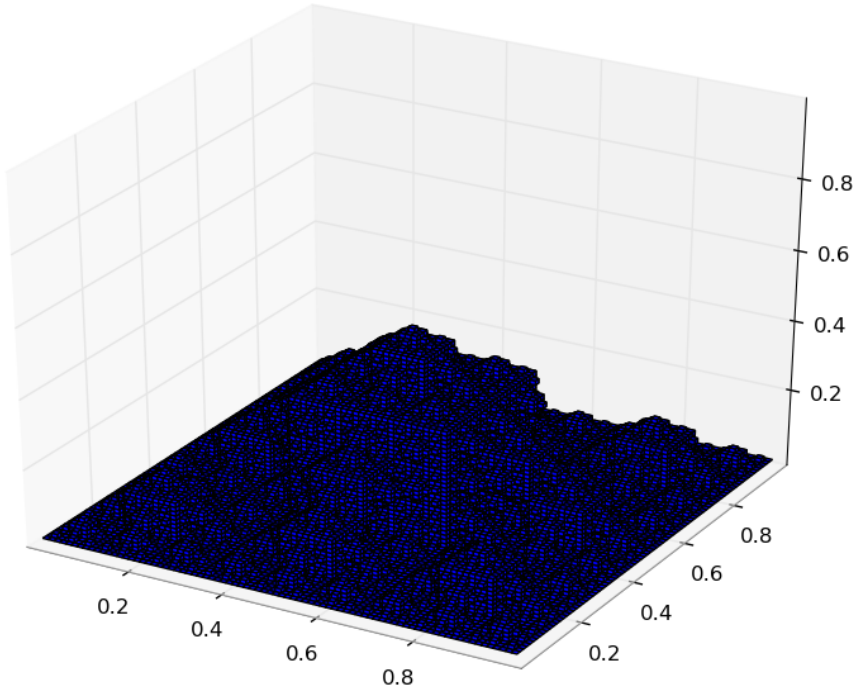

Figure 2: image of the fractal K4

Homework 2

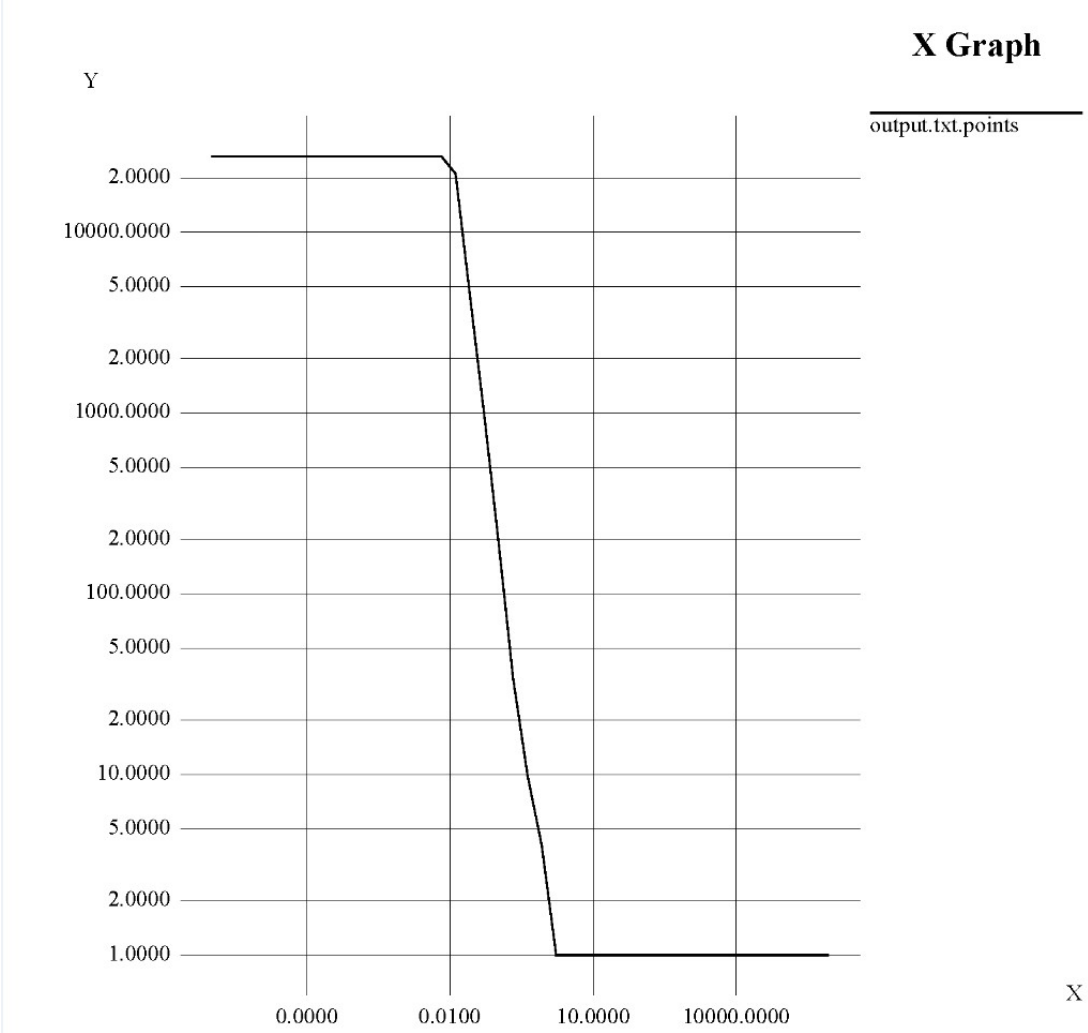

Figure 3: Hausdorff plot for K4

Hausdorff fractal dimension of K4: 2.07995

Homework 2

Q3:

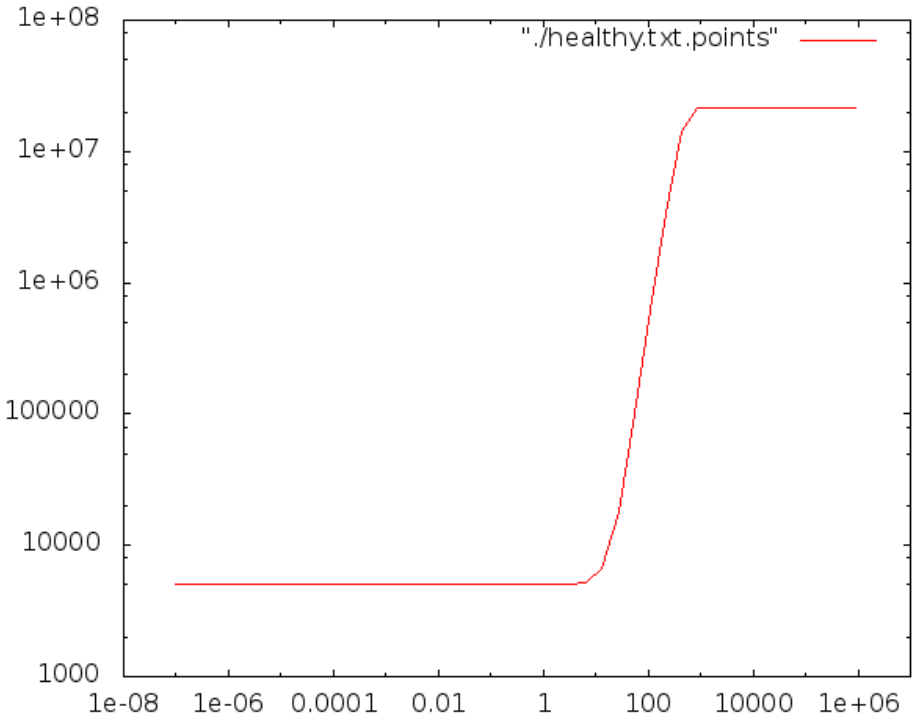

Figure 4:
Correlation plot of dataset H
Correlation Integral: 2.52976

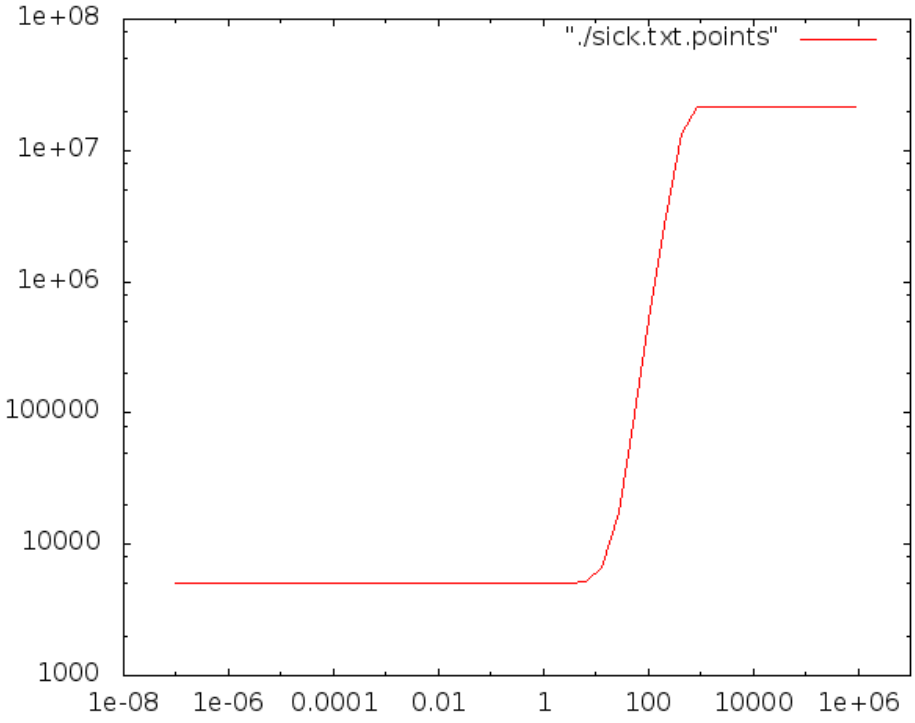

Figure 5:
Correlation plot of dataset S
Correlation Integral: 2.51242

Homework 2

a) **Is dataset H uniformly distributed in 4D space?**
   No. Dataset H is uniform distribution only if correlation integral is 4.

b) **Is dataset S uniformly distributed in 4D space?**
   No. Dataset S is uniform distribution only if correlation integral is 4.

Are H and S separable? No.

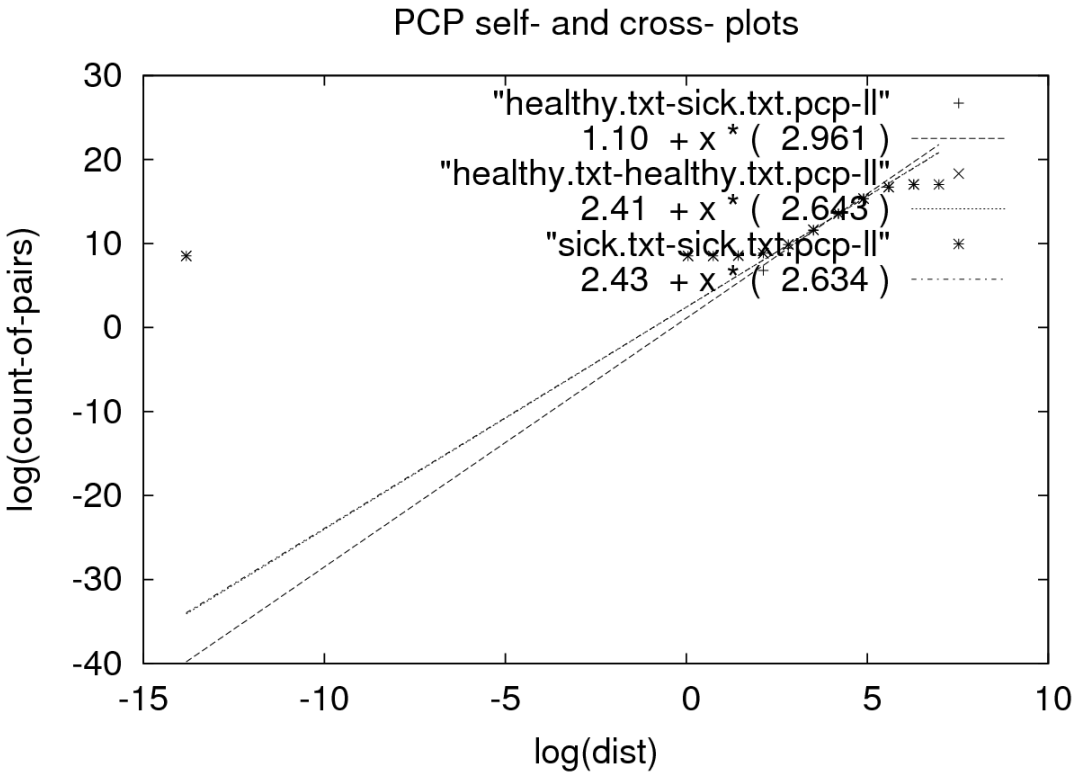

Figure 6:
PCP self- and cross- plots between healthy and sick dataset

Justification:

Tool used: BOPS plot
Through the BOPS plot in Figure 6, we can see that self- plots of healthy and sick dataset are overlapped and can't be visually separated (should have three lines in the graph, two self- plots are overlapped and the other is cross- plot). In addition, as we can see from the plot of 3D projection of the 4D patient dataset in the handout, sick and healthy points are already impossible to separate. Under this circumstance, given that the correlation dimension of both data sets is no greater than 3 and almost identical, it is not even possible to separate them in 4D plot.

Homework 2

Q4:

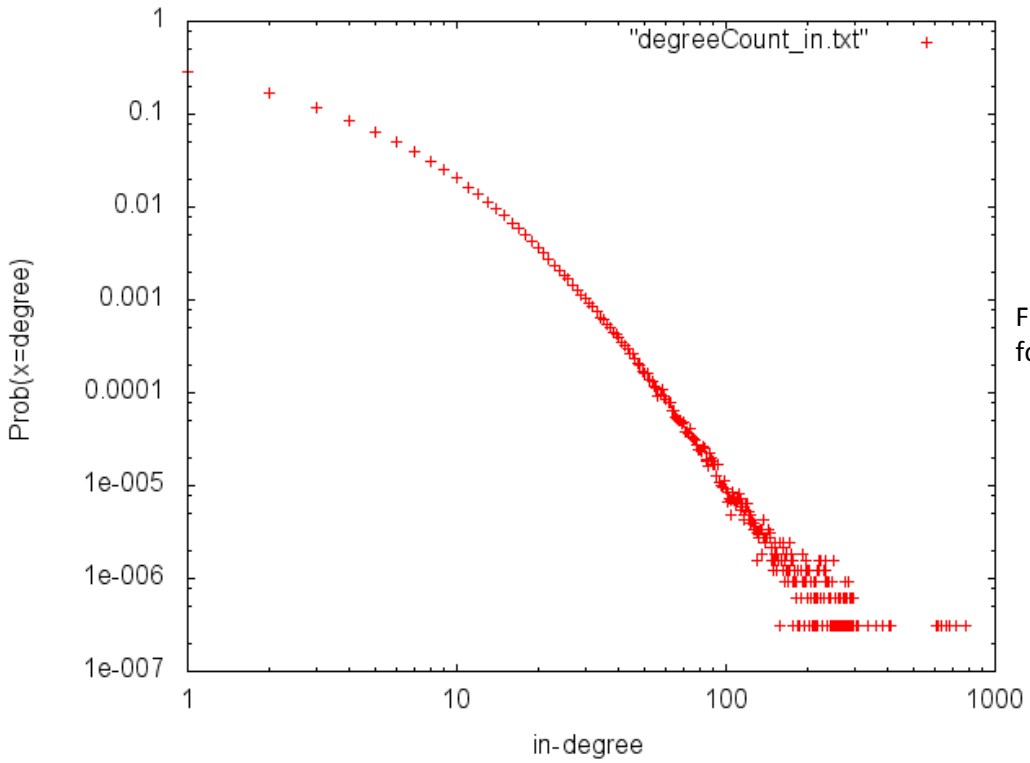

Figure 7: PDF plot for in-degree

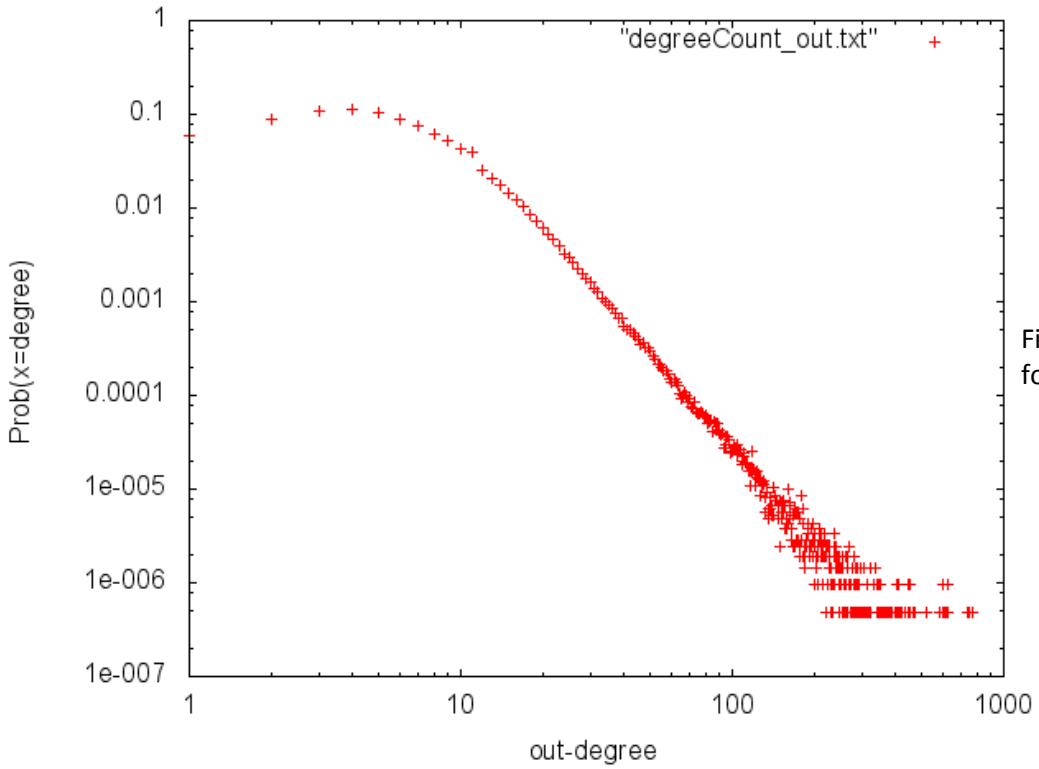

Figure 8: PDF plot for out-degree

Homework 2

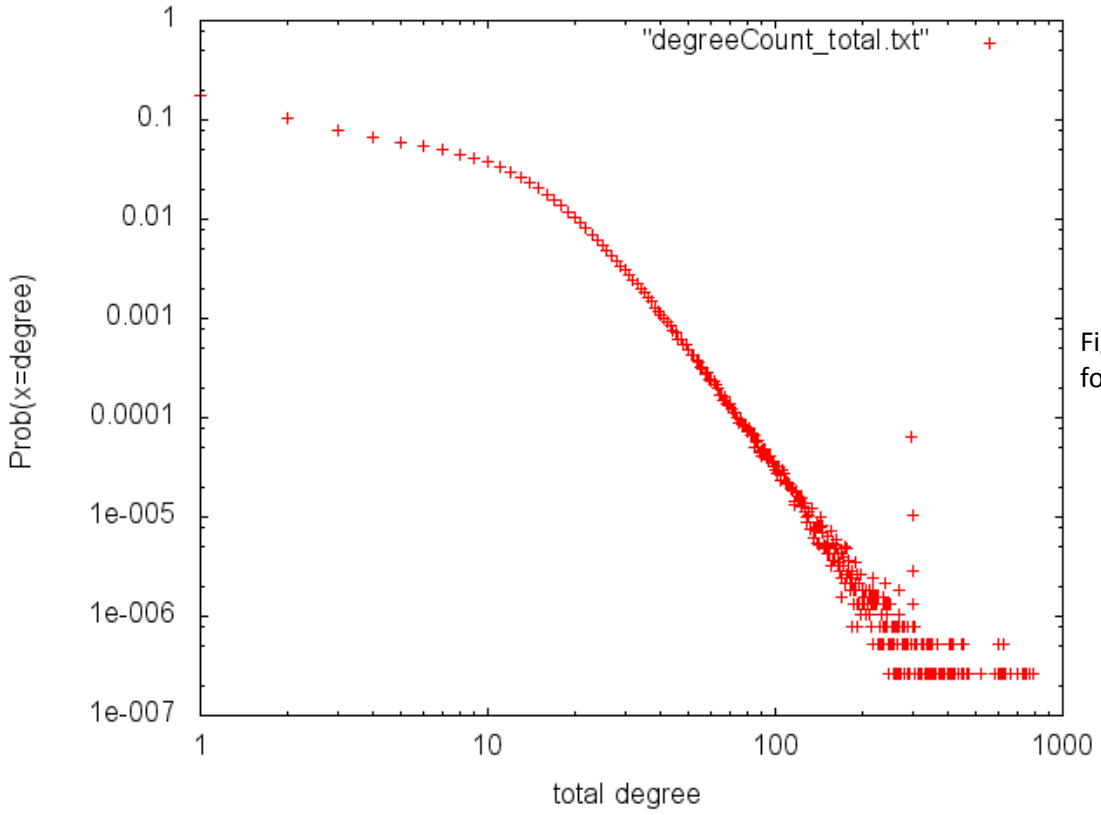

Figure 9: PDF plot for total-degree

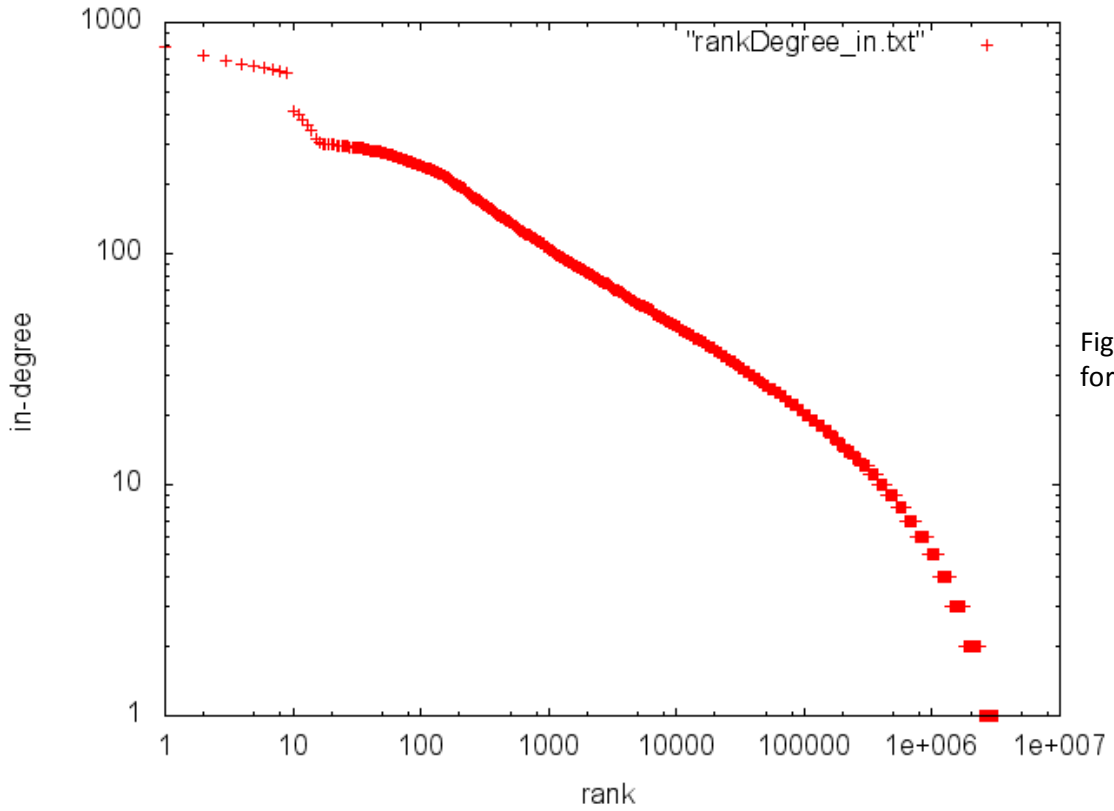

Figure 10: Zipf plot for in-degree

Homework 2

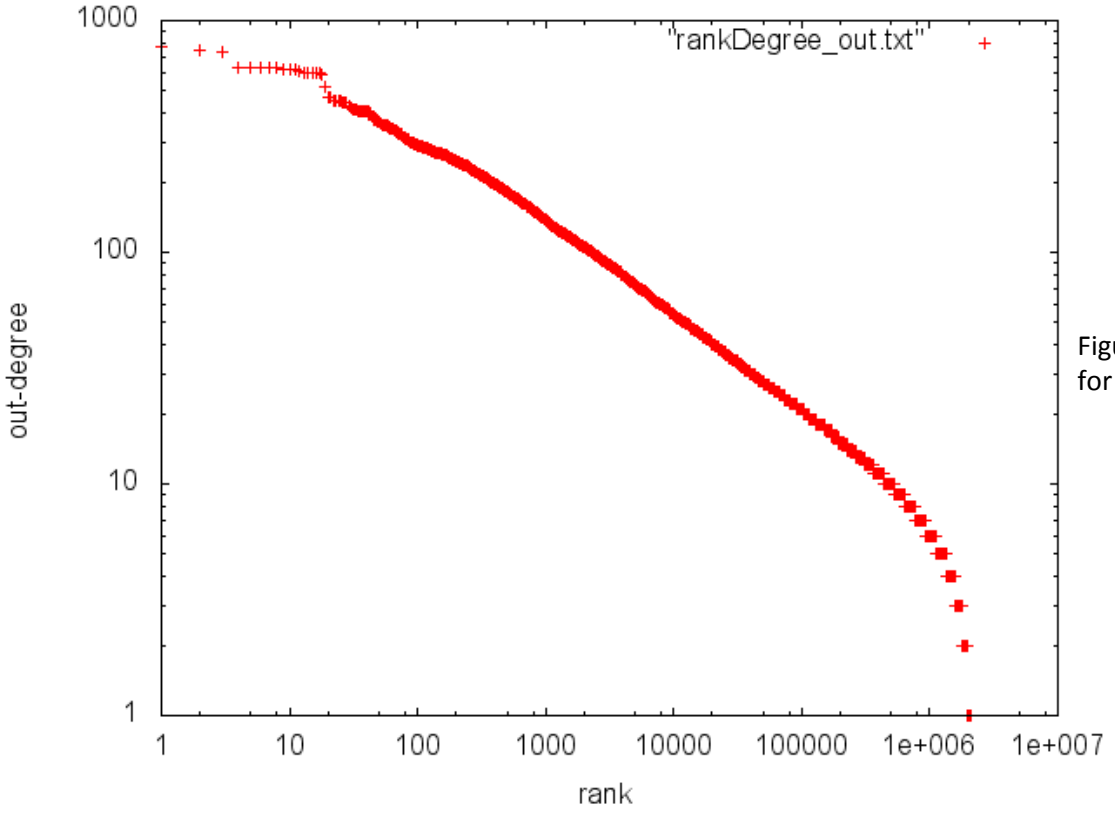

Figure 11: Zipf plot for out-degree

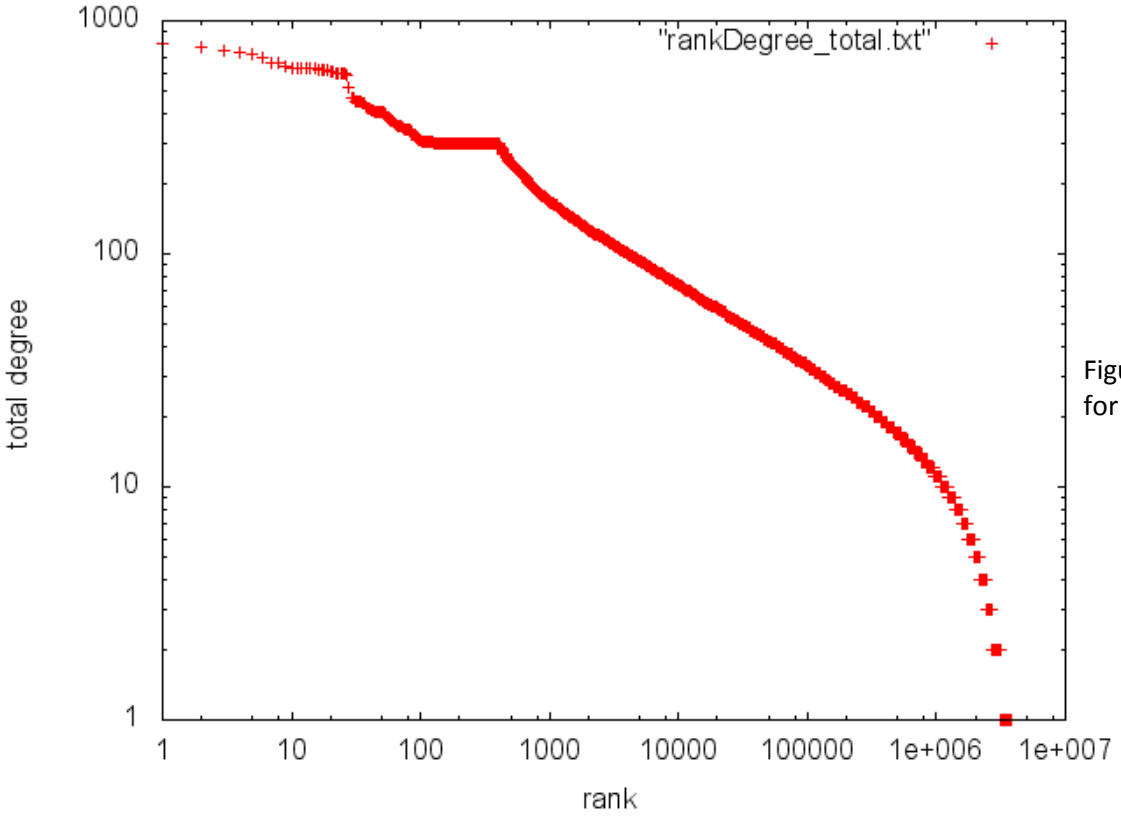

Figure 12: Zipf plot for total-degree

Homework 2

Estimation for the size K: 300

Justification:

Because as we can see in the Zipf plot for total-degree, there is a plateau at ranks 100-through-400, which could be due to a clique of size 300.