Principles of Complex Systems, CSYS/MATH 300
University of Vermont, Fall 2015
Assignment 7 • code name: Slice of heaven

Dispersed: Saturday, November 7, 2015.
Due: Friday, November 13, 2015.
Some useful reminders:
Deliverator: Peter Dodds
Office: Farrell Hall, second floor, Trinity Campus
E-mail: peter.dodds@uvm.edu
Office hours: 10:30 am to 11:30 am on Mondays, and 2:00 pm to 3:30 pm Wednesdays at Farrell Hall
Course website: http://www.uvm.edu/~pdodds/teaching/courses/2015-08UVM-300

All parts are worth 3 points unless marked otherwise. Please show all your workingses clearly and list the names of others with whom you collaborated.

Graduate students are requested to use LaTeX (or related TeX variant).

Email submission: PDF only! Please name your file as follows (where the number is to be padded by a 0 if less than 10 and names are all lowercase):
`CSYS300assignment%02d$firstname-$lastname.pdf` as in
`CSYS300assignment06michael-palin.pdf`

Please submit your project's current draft in pdf format via email. Please use this file name format (all lowercase after CSYS):
`CSYS300project-$firstname-$lastname-YYYY-MM-DD.pdf` as in
`CSYS300project-lisa-simpson-1989-12-17.pdf`

1. (3 + 3 + 3 + 3) This question is all about pure finite and infinite random networks

   We'll define a finite random network as follows. Take N labelled nodes and add links between each pair of nodes with probability p.

   (a) i. For a random node i, determine the probability distribution for its number of friends k, $P_k(p, N)$.

   ii. What kind of distribution is this?

   iii. What does this distribution tend toward in the limit of large N, if p is fixed?
   (No need to do calculations here; just invoke the right Rule of the Universe.)

(b) Using $P_k(p, N)$, determine the average degree. Does your answer seem right intuitively?

(c) Show that in the limit of $N \to \infty$ but with mean held constant, we obtain a Poisson degree distribution.

Hint: to keep the mean constant, you will need to change p.

(d) i. Compute the clustering coefficients C_1 and C_2 for standard finite random networks (N nodes).

ii. Explain how your answers make sense.

iii. What happens in the limit of an infinite random network with finite mean?

2. (3 + 3)

More on the peculiar nature of distributions of power law tails:

Consider a set of N samples, randomly chosen according to the probability distribution $P_k = ck^{-\gamma}$ where $k \geq 1$ and $2 < \gamma < 3$. (Note that k is discrete rather than continuous.)

(a) Estimate $\min k_{\max}$, the approximate minimum of the largest sample in the network, finding how it depends on N.

(Hint: we expect on the order of 1 of the N samples to have a value of $\min k_{\max}$ or greater.)

Hint—Some visual help on setting this problem up:

Direct link: `http://www.youtube.com/v/4tqlEuXA7QQ?rel=0`

(b) Determine the average value of samples with value $k \geq \min k_{\max}$ to find how the expected value of $k_{\max}$ (i.e., $\langle k_{\max} \rangle$) scales with N.

For language, this scaling is known as Heap's law.

3. (3 + 3)

   Let's see how well your answer for the previous question works.

   For $\gamma = 5/2$, generate $n = 1000$ sets each of $N = 10$, 10^2, 10^3, 10^4, 10^5, and 10^6 samples, using $P_k = ck^{-5/2}$ with $k = 1, 2, 3, \ldots$

   How do we computationally sample from a discrete probability distribution?

   Hint: You can use a continuum approximation to speeds things up.

   (a) For each value of sample size N, plot the maximum value of the $n = 1000$ samples as a function of sample number (which is not the sample size N). So you should have $k_{\max}$ for $i = 1, 2, \ldots, n$ where i is sample number. These plots should give a sense of the unevenness of the maximum value of k, a feature of power-law size distributions.

   (b) For each set, find the maximum value. Then find the average maximum value for each N. Plot $\langle k_{\max} \rangle$ as a function of N and calculate the scaling using least squares.

       Does your scaling match up with your theoretical estimate?