1. (3 + 3):

Consider a modified version of the Barabási-Albert (BA) model where two possible mechanisms are now in play. As in the original model, start with \( m_0 \) nodes at time \( t = 0 \). Let’s make these initial guys connected such that each has degree 1. The two mechanisms are:

M1: With probability \( p \), a new node of degree 1 is added to the network. At time \( t + 1 \), a node connects to an existing node \( j \) with probability

\[
P(\text{connect to node } j) = \frac{k_j}{\sum_{i=1}^{N(t)} k_i}
\]  

where \( k_j \) is the degree of node \( j \) and \( N(t) \) is the number of nodes in the system at time \( t \).

M2: With probability \( q = 1 - p \), a randomly chosen node adds a new edge, connecting to node \( j \) with the same preferential attachment probability as above.

Note that in the limit \( q = 0 \), we retrieve the original BA model (with the difference that we are adding one link at a time rather than \( m \) here).

In the long time limit \( t \to \infty \), what is the expected form of the degree distribution \( P_k \)?

Do we move out of the original model’s universality class?
Different analytic approaches are possible including a modification of the BA paper, or a Simon-like one (see also Krapivsky and Redner [2]).

Hint: You can attempt to solve the problem exactly and you’ll find an integrating factor story.

Another hint, moment of mercy: Approximate the differential equation by considering large $t$ (this will simplify the denominators).

(3 points for set up, 3 for solving.)

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 $P_k = c k^{-\gamma}$ where $k \geq 1$ and $2 < \gamma < 3$. (Note that $k$ is discrete rather than continuous.)

(a) Estimate $\min k_{\text{max}}$, 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 $\min k_{\text{max}}$ or greater.)

Hint—Some visual help on setting this problem up:
http://www.youtube.com/v/4tqlEuXA7QQ?rel=0

(b) Determine the average value of samples with value $k \geq \min k_{\text{max}}$ to find how the expected value of $k_{\text{max}}$ (i.e., $\langle k_{\text{max}} \rangle$) 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 $N = 10, 10^2, 10^3, 10^4, 10^5, \text{ and } 10^6$ samples, using $P_k = c k^{-5/2}$ with $k = 1, 2, 3, \ldots$

Question: how do we computationally sample from a discrete probability distribution? Does our approach differ for large and small spaces?

Hint: A continuum approximation will help.

(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 $k_{\text{max}}$ 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 $\langle k_{\text{max}} \rangle$ as a function of $N$ and calculate the scaling using least squares.
Does your scaling match up with your theoretical estimate?

References
