Theory and Applications of Complex Networks

Class Four

College of the Atlantic

David P. Feldman

23 September 2008

`http://hornacek.coa.edu/dave/`

1. A Statistical Tangent
2. Recap of Erdős-Rényi model
3. Description of Small-World model
4. Properties of the Small-World model

Null Models and p-values

- At a certain school, two thirds of the students are women.
- A certain class of 8 students has 7 women in it.
- Is this unusual?
- Or is this something that could happen by chance?

- **Null Hypothesis:** Men and women are equally likely to take the class.
- **Alternative Hypothesis:** Men and women are not equally likely to take the class.

Can we Reject the Null?

- Under the null hypothesis, the probability that there are k women in a class of N is given by:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \ . \qquad (1)$$

- For $N = 8$, how likely is it that there are 7 or more women in the class?

$$P(k \geq 7) = P(7) + P(8) = 0.1561 + 0.039 = 0.1952 \ . \qquad (2)$$

- The quantity 0.1562 is known as the p-value.
- The p-value is defined to be the probability that the null model would generate a result at least as extreme as the one which was actually observed.
- The experimenter sets a significance level α, often 5%.
- In this case, there is not significant evidence to reject the null.
- The smaler the p-value, the more evidence there is against the null.

Calculating p-values

- Three options:
  1. Look them up in a table
  2. Calculate by hand
  3. Simulate
- I wrote a short program to simulate choosing 8 students where each student is female with probability $\frac{2}{3}$.
- Running the simulation 10,000 times I get $p = 0.1873$.
- Running the simulation 100,000 times I get $p = 0.19622$.

The Erdős Rényi Model

1. Start with N nodes.
2. Connect each pair of nodes with probability p.

- The mean degree is $z = Np$
- Note that there are a number of different ways to consider the large N limit.
- Often, we want N to get large while keeping z constant.
- In science, we frequently need to ask, Could this have happened randomly, by chance?
- In order to answer this question, we need to know about random graphs.

Summary of Properties of Erdős-Rényi Model

- Degree distribution is Poisson:

$$P(k) = \frac{z^k e^{-z}}{k!} . \quad (3)$$

- Very low clustering:

$$C = \frac{z}{N} . \quad (4)$$

- Highly connected, “Small-world”:

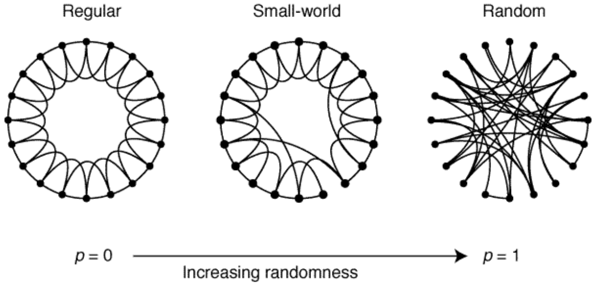
$$\ell \approx \log N . \quad (5)$$

- Connectivity properties change discontinuously as p is varied.

The Small-World Model

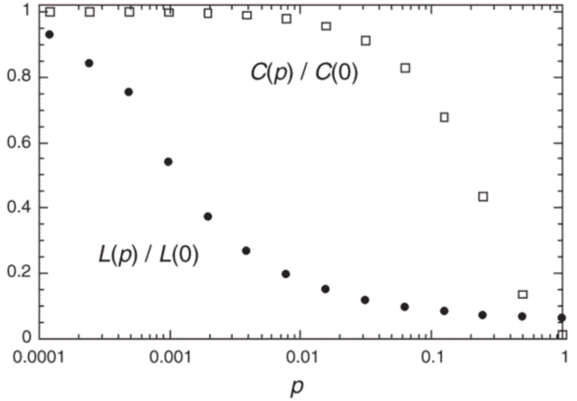
- The model:
  1. Begin with a regular lattice. Usually this is a one-dimensional ring, where each node has a few neighbors.
  2. Go through the regular lattice and consider each link.
  3. With probability p, rewire the link by random rewiring
- Initial question:
  1. How do C and ℓ vary with p?
- Watts and Strogatz, Nature 393:440–2. 1998.
- See also Newman, Models of the Small World,” Journal of Statistical Physics 101:819-841. 2000.

Watts-Strogatz Model



- As p is increased the model moves from a regular graph, through intermediate graphs, to a random graph at $p = 1$.
- Figure source `http://www.nature.com/nature/journal/v393/n6684/fig_tab/393440a0_F1.html`

Watts-Strogatz Model: Basic Results



- There is a large intermediate region which shows "small-world" behavior: small ℓ but large C.
- Note the log scale on the horizontal axis.
- Figure source `http://www.nature.com/nature/journal/v393/n6684/fig_tab/393440a0_F2.html`

Watts-Strogatz: Preliminary Conclusions

1. The WS model shows a transition from a large-world to a small-world.
2. Disease models which have a non-automated susceptibility to infection exhibit a sharp transition between epidemic and non-epidemic behavior.
3. Dynamical systems on small-world graphs exhibit behavior which is qualitatively different from behavior on regular graphs.
4. Many graphs show additional features (e.g., long-tailed degree distributions) which are not accounted for by the WS and similar models.
5. Nevertheless, the WS model qualitatively captures the small-world feature of many networks, and is a useful, albeit quite basic, model for a social network.

- Adapted from conclusions in Newman's 2003 review article.

Questions

- How do small-world networks grow?
- What sort of models might give us insight into networks in which the degree distribution is long-tailed?
- When are small-world networks navigable with local information?
- How does the behavior of dynamical systems (e.g., epidemic models or scheduling problems) change as network topology changes?
- How robust are results based on the WS model?