Solution 7-6

Ex 7.6: Loners, pairs and triads [3]

A sailing club has $N = 1000$ members. Friendships in the club form an ER graph with mean degree $z = 4$.

a) Calculate the number of isolated nodes, i.e. the number of members who have no friends in the club. (Hint: $0! = 1$)

Solution

The probability that a randomly picked node has no friends in the club is p_0. To get the total number of such nodes we multiply p_0 with the total number of members N. Using the Poisson distribution we find

$$Np_0 = N\frac{z^0 e^{-z}}{0!} = Ne^{-z} = 1000e^{-4} \approx 18 \tag{1}$$

b) Calculate the number of isolated pairs of friends in the club, i.e. the number of components of size 2. (Hint: This is not the same as the number of link motifs, because motifs are not necessarily isolated. You will have to find a different way.)

Solution

There are different ways to find the answer. For example we can think of an isolated pair as a node of degree 1 that links to a node of excess degree 0.

Using the degree distribution we can compute the number of nodes of degree 1

$$Np_1 = N\frac{z^1 e^{-z}}{1!} = 1000\frac{4^1 e^{-4}}{1!} = 4000e^{-4} \tag{2}$$

Furthermore we know from a previous exercise $p_k = q_k$ so the probability that the outgoing link from a node of degree 1 leads to a node of excess degree 0 is

$$q_0 = p_0 = \frac{z^0 e^{-4}}{0!} = e^{-4} \tag{3}$$

We can find the number of isolated pairs by multiplying the number of nodes of degree 1 with the probability that a link leads to a node of excess degree 0 and dividing by 2 to correct the double-counting that occurs because we are counting the link from both sides. The result is

$$\frac{Np_1 q_0}{2} = 2000e^{-8} \approx 0.7 \tag{4}$$

The expected number of isolated pairs ca. 0.7 so it seems unlikely that there is more than one such isolated pair.

An alternative way of computing this number is to consider that there are $N(N-1)/2$ ways in which we can pick two nodes from the system. To be considered an isolated pair

the two nodes must be linked (probability p). Each of the two nodes has $N - 2$ chances to receive a link to a node outside the pair. For an isolated pair all of these trials must fail (probability $q = 1 - p$ per such link), hence we can compute the expected number of isolated pairs as

$$n_{\text{isolater pair}} = \frac{N(N - 1)}{2} pq^{2(N-2)} \tag{5}$$

where we have multiplied the number of pairs with the probability that the link connecting the pair is present and the probability that all links that would connect the pair to another node are absent. We know $N = 1000$ and $z = 4$ and hence

$$p = \frac{4}{999} \qquad q = \frac{995}{999} \tag{6}$$

Taking these values into account we find

$$n_{\text{isolater pair}} = \frac{1000 \cdot 999}{2} \frac{4}{999} \left(\frac{995}{999}\right)^{1996} \approx 0.7 \tag{7}$$

as before.

c) Finally compute the expected number of components of size 3.

Solution

A component of size 3 could be an isolated triangle, or an isolated 3-chain (o-o-o).
   You might simply guess that isolated triangles are super rare in any ER-graph ($N = 3$, $p = 1$ is perhaps the only exception to this rule)
   One way to convince ourselves of this is to use our equation for the number of triangles

$$N_\Delta = \frac{1}{6} z^3 = \frac{64}{6} = \frac{32}{3}, \tag{8}$$

so only about 10. But this is the total number of triangles, not the number of isolated triangles. During the network creation each of the three nodes in these triangles could have become connected to each of the 997 external nodes. So there were $3 * 997 = 2991$ chances to connect any given triangle to at least one additional node making it non-isolated. For an isolated triangle we want all of these trials to fail. For one trial the chance of failure is $q = 995/999$ as we've already computed in the previous part. So the expected number of triangles is

$$n_{\text{isolated } \Delta} = \frac{64}{6} \left(\frac{995}{999}\right)^{2991} \approx 0.00007 \tag{9}$$

so indeed the chances of seeing even just one isolated triangle are very very slim.
   So how about the isolated 3-chain? A quick way to compute their number is to first compute the number of nodes of degree 2

$$n_2 = N \frac{z^2 e^{-z}}{2!} = 8000 e^{-4} \approx 146.5 \tag{10}$$

The probability that a random link leads to a node of degree 1 (i.e. a node with no additional links) is q_0. We know from above

$$q_0 = p_0 = e^{-4} \tag{11}$$

2

The probability that both nodes that start from a node of degree 2 lead to a node of degree 1 is $q_0{}^2 = e^{-8}$, so the expected number of nodes of degree 2 in our network is

$$n_{\text{isolated 3-chain}} = 8000e^{-12} \approx 0.05 \tag{12}$$

This is much more likely than finding an isolated triangle, but still so unlikely that there's only a ca. 5% chance that our sailing club contains an isolated 3-chain. [Note this is an expectation value not a probability, but if the expectation value is so small the chance that a nework has two or more copies of the motif is practically zero. In this case our result is literally saying that ca. 5% of random networks will have one isolated 3-chain and the other 95% have none.]