COT 6938 Network Science Assignment 1

Heather Lawrence

February 5, 2016

Q1

Write a program in any language to: consider a square region of 100x100 units. Suppose only 100 nodes are *uniformly randomly* scattered over this square region. Do not confine the coordinates to integers. Two nodes are connected if their distance is less than d units. Find the adjacency matrix for the network when d = 20. i) Find and plot the degree distribution for d = 20. Use the same coordinates of the 100 nodes for all three cases. Keep the same x and y-axis for all three plots. ii) Find and plot the degree distribution for d = 30. iii) Find and plot the degree distribution for d = 40.

As the Euclidean distance where an edge will form between two nodes is increased the number of edges, and thus the number of connections a node has to another node, increases. In 1 the distance between two nodes where an edge will form is 20 units. As the distance increases in 2 and 3 the number of edges formed increases. This is reflected in the degree distribution observed as the distance increases in 4. When d = 20 the nodes have fewer possible neighbors causing the 99 nodes used to show fewer degrees. When d = 40 the nodes have more possible neighbors causing the 99 nodes to show more degrees. With a higher range of possible neighbors the probability a specific degree occurs is less (the curve flattens).

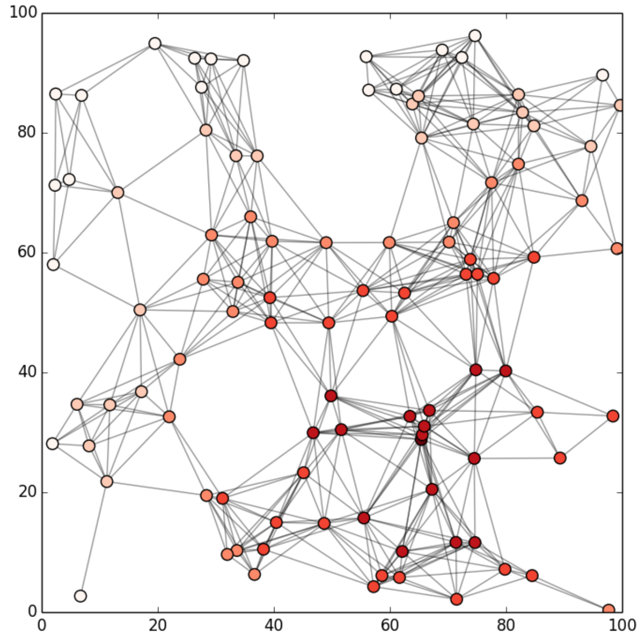

Figure 1: Random Graph Over 100x100 when d = 20

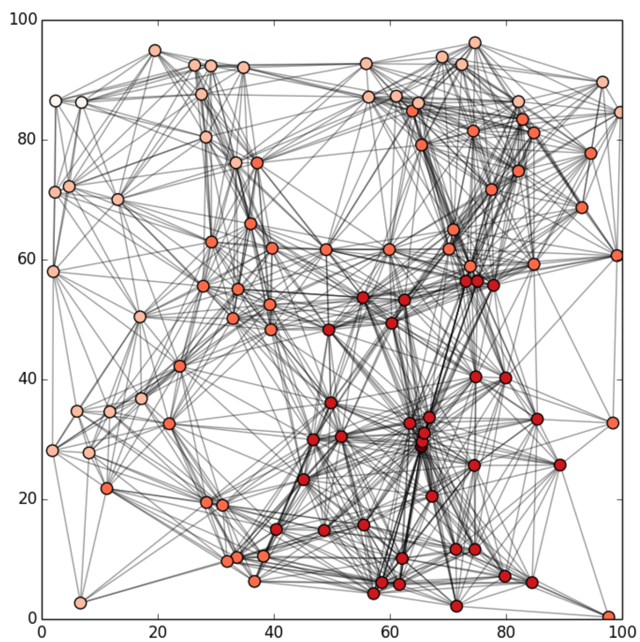

Figure 2: Random Graph Over 100x100 when d = 30

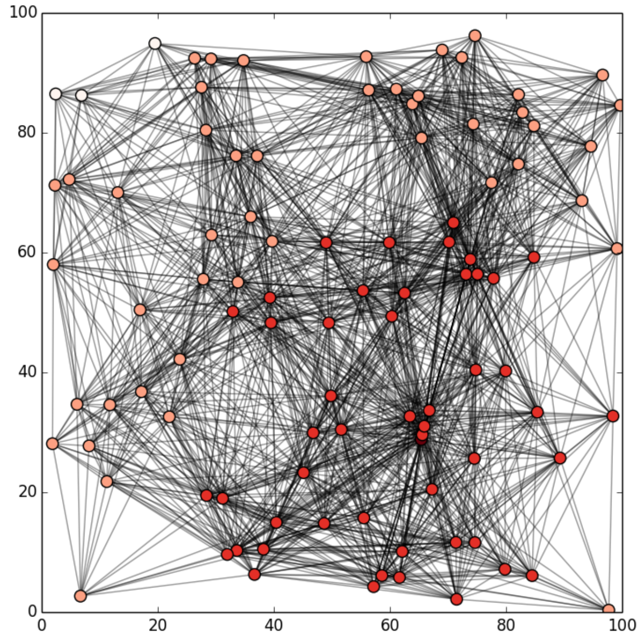

Figure 3: Random Graph Over 100x100 when d = 40

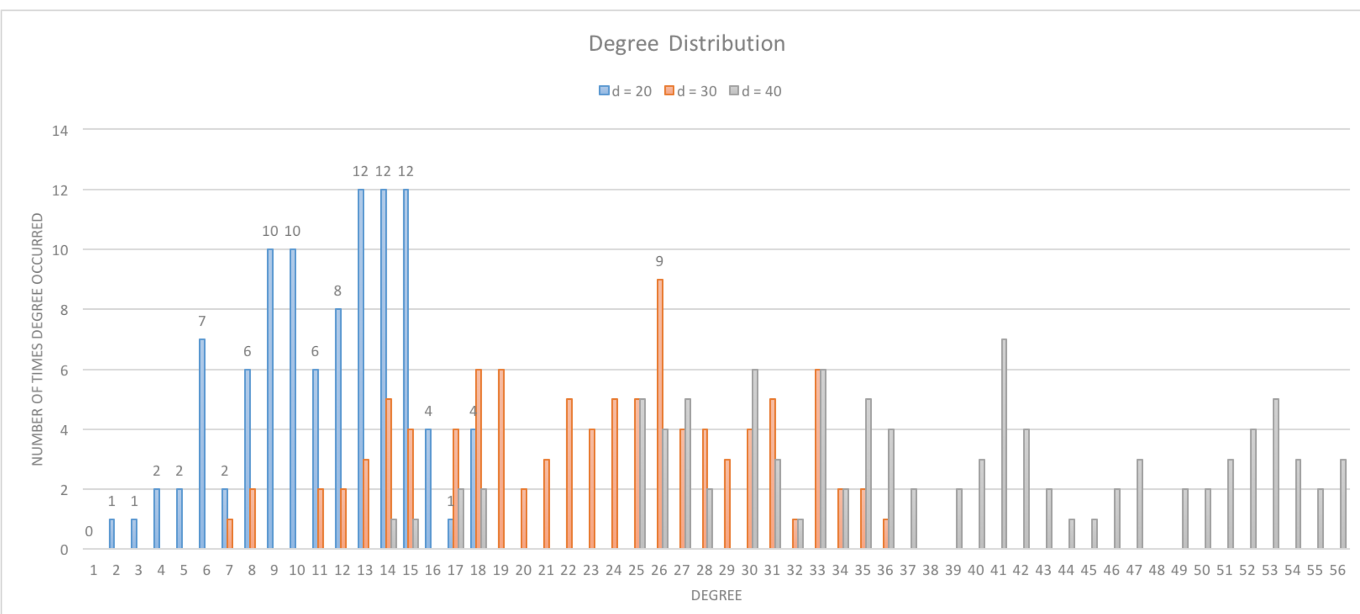

Figure 4: Degree Observed vs. Number of Times Occurred

Q2

Assume the same set up as in Q1, 100 nodes scattered over a 100x100 square grid. Show all steps you used to compute the probabilities. Submit the plot. Use Poisson distribution to find the following (do not use Binomial distribution):

i) What is the probability that a node located somewhere at the center of a square region has no links with the other 99 nodes when d = 30. You can use the calculator or write a program to find this probability.

Plotted a new random network on the 100x100 field indicated in figure 5. Node 43: (41.286797855396074, 45.74515750808099) is located near the center of the field.

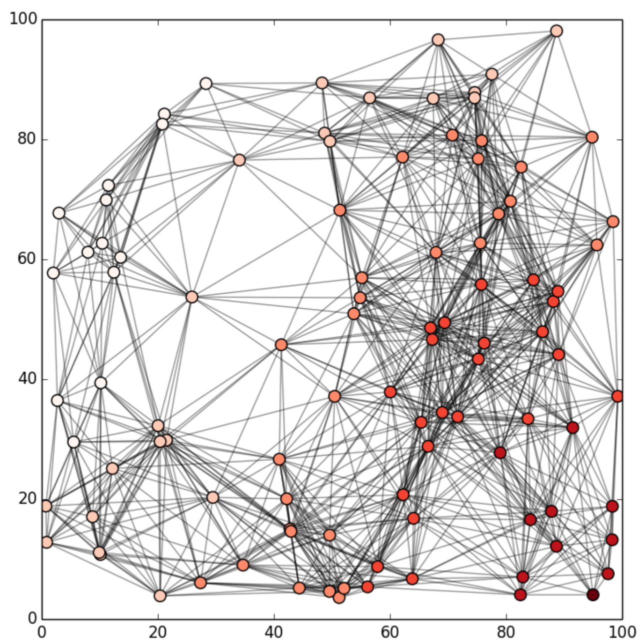

Figure 5: Random Graph Over 100x100 when d = 30

A link generates between two nodes in the event the two nodes are within d = 30. In order for a node to generate no links there must be no nodes in the area.

$$p\{\text{no nodes fall within d}\} = \{\text{probability all nodes fall outside d}\} =$$

$$\left(\frac{A-\pi r^2}{A}\right)^{99} = \left(\frac{10000-900\pi}{10000}\right)^{99} = 0.717256^{99} = \boxed{5.1497e^{-15}}$$

ii) What is the probability that the same node has exactly 1 link when d = 30?

$$\text{P}\{\text{single node falls within d}\} = \frac{\pi(r^2)}{A} = \frac{900\pi}{10000} = \boxed{0.282743}$$

iii) What is the probability that the node has exactly 2 links when d = 30?
P{2 nodes falls within d} = 0.282743 * 0.282743 = $\boxed{0.079944}$

iv) repeat this process for 3, 4, 5, 6... links. Essentially what you obtain is the degree distribution. Just use a 'for' loop. v) What is the expected degree of that node?

if p{1} = 0.282743,

$$\text{the average degree } < k >= p(n-1) = 0.282743(99-1) = 27.708 \approx 28$$

vi) Plot the degree distribution (the number of links in the x-axis and corresponding probabilities on the y-axis)

(0, 7.180091001121677e-15) effectively 0
(1, 2.7737868860555473e-13) effectively 0
(2, 5.303125384334231e-12) effectively 0
(3, 6.689581360289272e-11) effectively 0
(4, 6.262959746034742e-10) effectively 0
(5, 4.641459186385674e-09) effectively 0
(6, 2.8359808922620198e-08) effectively 0
(7, 1.469298902001799e-07) effectively 0
(8, 6.588379780509499e-07) effectively 0
(9, 2.597141978789468e-06) effectively 0
(10, 9.11177034051804e-06) effectively 0
(11, 2.8734908479265787e-05) effectively 0
(12, 8.212301609340806e-05) effectively 0
(13, 0.00021415935804789521)
(14, 0.0005125603560972628)
(15, 0.0011314884002225798)
(16, 0.0023137971513142283)
(17, 0.004399540088734393)
(18, 0.007804352184629257)
(19, 0.012953594918527728)
(20, 0.020169933842116117)
(21, 0.029532271200730674)
(22, 0.04074571645362635)
(23, 0.05307436186997817)
(24, 0.06538104672646641)
(25, 0.07628877592266177)
(26, 0.08443591713672766)
(27, 0.08875912731973724) highest probability as expected
(28, 0.08872181287476741) highest probability as expected
(29, 0.08442042140310721)
(30, 0.07654069793698925)
(31, 0.06618456262436884)

(32, 0.05462589855940882)
(33, 0.0430671026881102)
(34, 0.032456161951523906)
(35, 0.02339516743507248)
(36, 0.01613917284274078)
(37, 0.010660763387975319)
(38, 0.0067460818961750935)
(39, 0.004091243410274751)
(40, 0.002378835469626909)
(41, 0.0013265560620376279)
(42, 0.0007096892717997072)
(43, 0.00036433834220305313)
(44, 0.00017952790263726578)
(45, 8.492396634740701e-05) effectively 0
(46, 3.8571397454321605e-05) effectively 0
(47, 1.682239885535011e-05) effectively 0
(48, 7.0458587222301725e-06) effectively 0

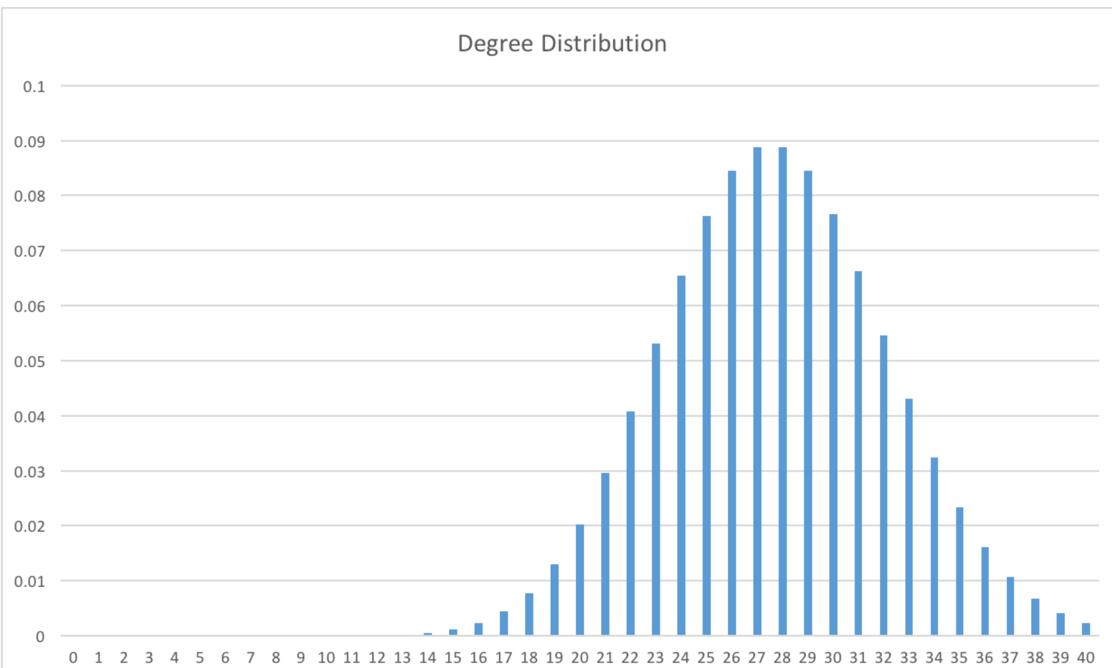

Figure 6: Calculated Degree Distribution

vii) Compare the plots you obtained for d = 30 from the program (Q1) and from the analysis. The best would be to draw on top of each other. Make sure you have the same range for the x and y axes. Do the plots agree with each other? Explain your answer

The numbers I obtained for Q2 are a normalized degree distribution and the numbers in Q1 indicate the times a specific degree occurred, graphing these on the same axes would not best represent both sets of numbers. A binomial distribution (bell curve) of Q1 and Q2 is apparent, but as both are random graphs (one generated, one calculated) they are not exactly the same.

Q3

Consider a random network with 10 nodes where a link between any two nodes exists with probability 0.3. What is the expected number of links?

N = 10
p = 0.3
L = ?

$$L = p\frac{N(N-1)}{2} = 0.3\frac{10(10-1)}{2} = 0.3\frac{90}{2} = 0.3(45) = 13.5$$

Can't have half a link, so $\boxed{L = \{12, 13\}}$

Find the probability that the network will have 30 links.
N = 10
L = 30
p = ?

$$p = \frac{L}{\frac{n(n-1)}{2}} = \frac{30}{\frac{10(10-1)}{2}} = \frac{30}{45} = \boxed{0.\overline{66}}$$

How many realizations of 30 links are possible?

Number of different ways to place L links among node pairs =

$$\binom{\frac{N(N-1)}{2}}{L} = \binom{45}{30} = \boxed{344,867,425,584}$$

Q4

Show that the local clustering coefficient of a node is independent of the node's degree in a random network.

$$C_i = \frac{\text{Number of connected neighbors}}{\text{all possible connections}}$$

$$e_i = p = \frac{k_i(k_i - 1)}{2}$$

$$C_i = \frac{e_i}{\frac{k_i(k_i-1)}{2}} = p\frac{\frac{k_i(k_i-1)}{2}}{\frac{k_i(k_i-1)}{2}}$$

$$C_i = p = \frac{< k >}{N}$$

Q5

Consider a large random network (i.e. use Poisson Approximation) whose average degree is 10.
i) What is the probability that a node chosen randomly has a degree more than 20?

$$P\{k > 20\} = \int_{k=20}^{\infty} \frac{e^{-10} * 10^k}{k!} dx = \int_{k=20}^{\infty} \frac{e^{-10} * 10^{20}}{20!} dx = \boxed{0.00158}$$

ii) What is the probability that a node chosen randomly has a degree less than 2?

$$P\{k < 2\} = \int_{k=2}^{\infty} \frac{e^{-10} * 10^k}{k!} dx = \int_{k=2}^{\infty} \frac{e^{-10} * 10^2}{2!} dx = \boxed{0.00049}$$

iii) What is the probability that a node chosen randomly has a degree between 8 and 12, inclusive?

Probability of ($8 \leq x \leq 12$) is the difference of the area under those two curves.

$$P\{k \geq 8\} = \int_{k=8}^{\infty} \frac{e^{-10} * 10^k}{k!} dx = \int_{k=8}^{\infty} \frac{e^{-10} * 10^8}{8!} dx = \boxed{0.112599}$$

$$P\{k \geq 12\} = \int_{k=12}^{\infty} \frac{e^{-10} * 10^k}{k!} dx = \int_{k=12}^{\infty} \frac{e^{-10} * 10^{12}}{12!} dx = \boxed{0.09478}$$

0.112599 - 0.09478 = $\boxed{0.17819}$