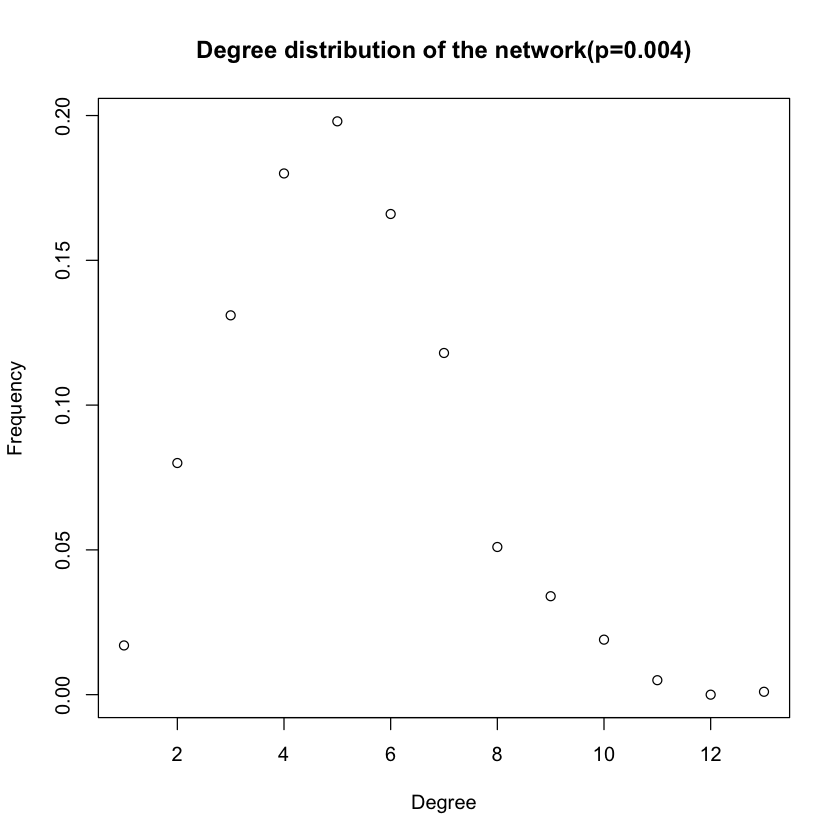
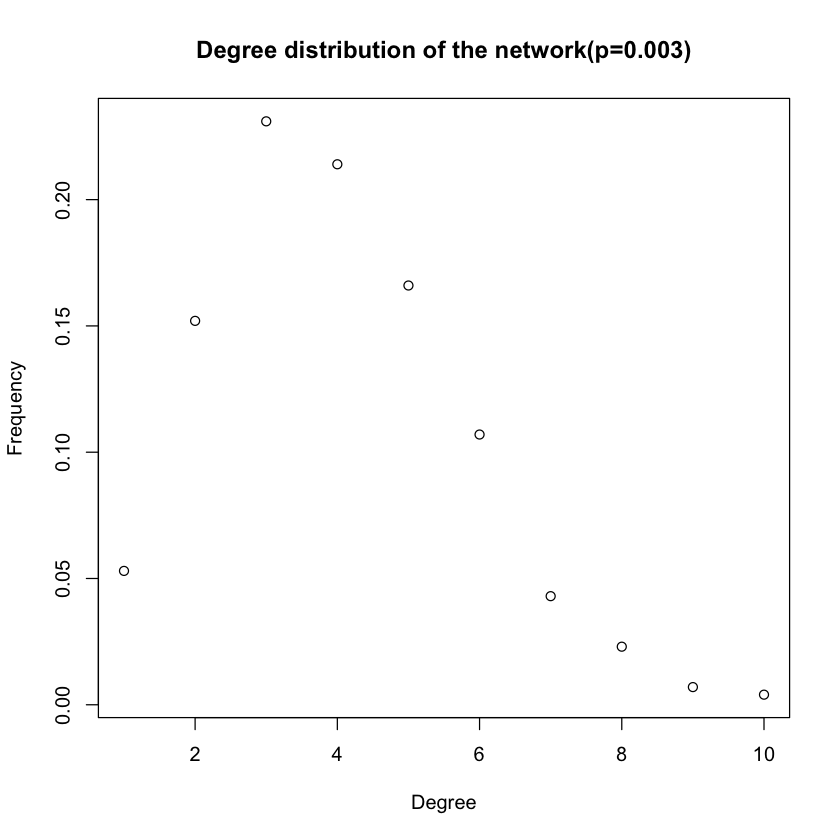
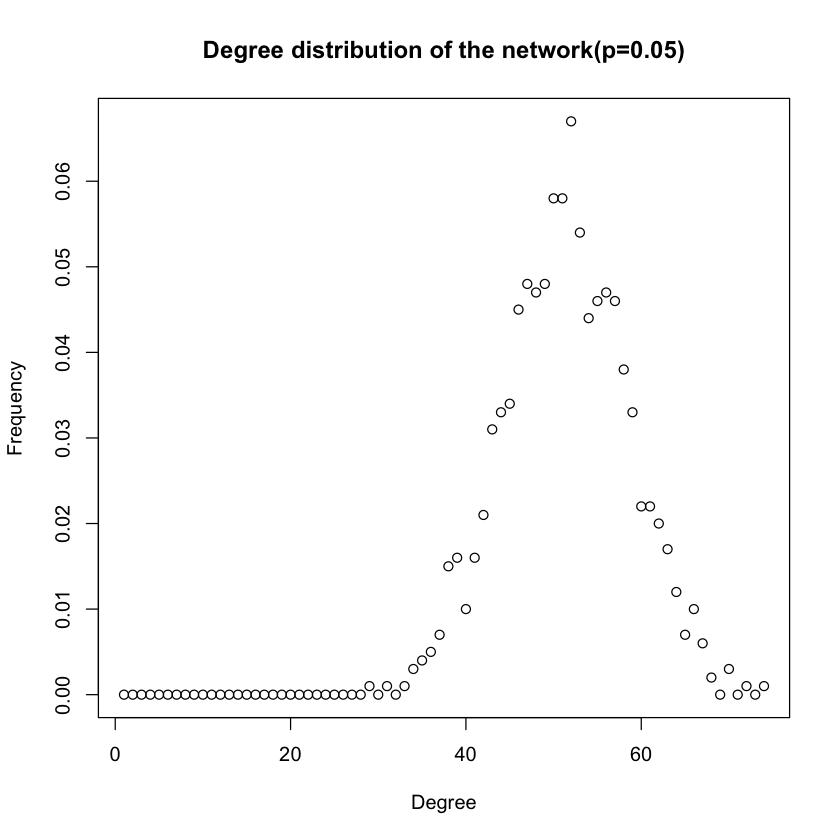
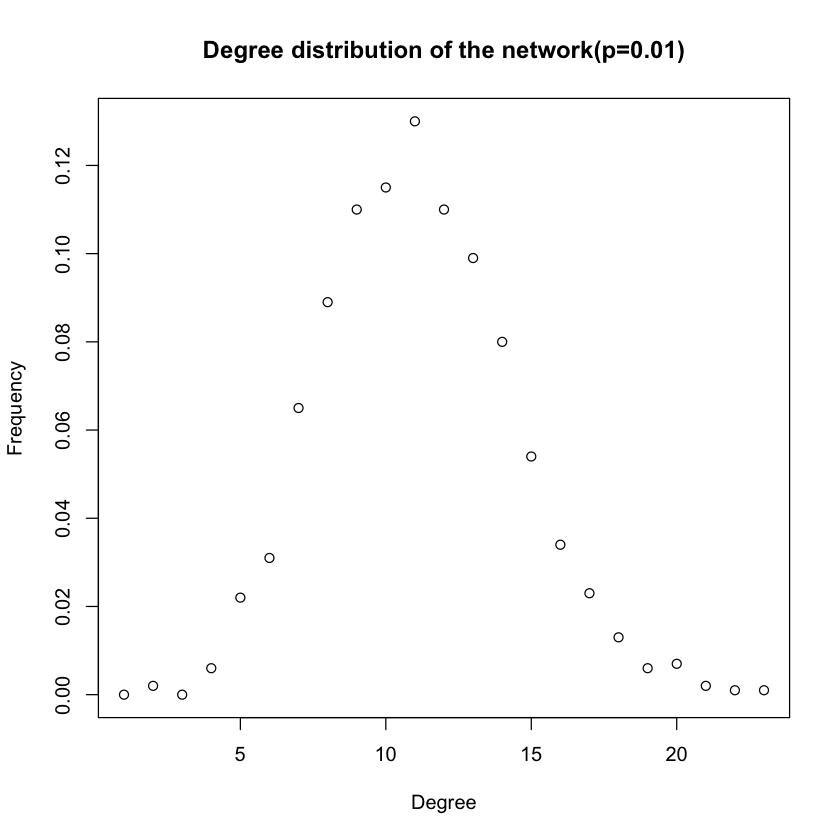
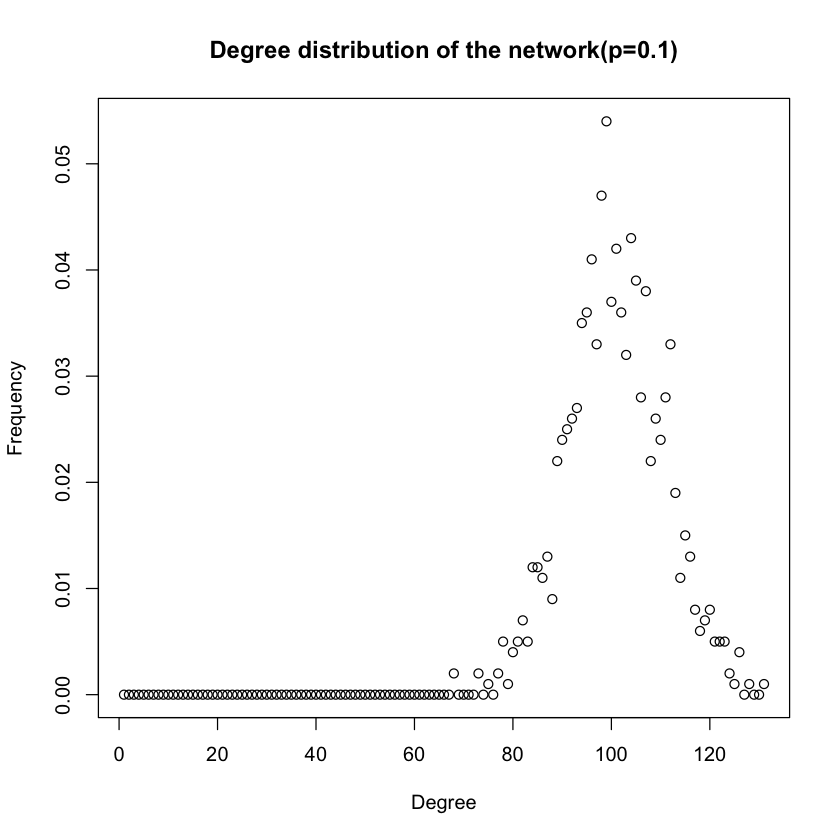
**1 Generating Random Networks**

1. Create random networks using Erdös-Rényi (ER) model
2. Create an undirected random networks with n = 1000 nodes, and the probability p for drawing an edge between two arbitrary vertices 0.003, 0.004, 0.01, 0.05, and 0.1. Plot the degree distributions. What distribution is observed? Explain why. Also, report the mean and variance of the degree distributions and compare them to the theoretical values.

We generate 5 graphs: g1(p=0.003), g2(p=0.004),g3(p=0.01)g4(p=0.05),g5(p=0.1)







-----g1-----

mean for g1 is 3.222

variance for g1 is 3.14386

-----g2-----

mean for g2 is 4.122

variance for g2 is 4.385502

-----g3-----

mean for g3 is 9.984

variance for g3 is 9.877622

-----g4-----

mean for g4 is 49.714

variance for g4 is 51.63384

-----g5-----

mean for g4 is 100.304

variance for g5 is 90.8164

theoritical value:

p=0.003: mean= np =3, variance is np(1-p)=2.991

p=0.004: mean= np =4, variance is np(1-p)=3.984

p=0.01: mean= np =10, variance is np(1-p)=9.9

p=0.05: mean= np =50, variance is np(1-p)=47.5

p=0.1: mean= np =100, variance is np(1-p)=90

(b) For each p and n = 1000, answer the following questions: Are all random realizations of the ER network connected? Numerically estimate the probability that a generated network is connected. For one instance of the networks with that p, find the giant connected component (GCC) if not connected. What is the diameter of the GCC?

g1 connectivity: FALSE

g2 connectivity: FALSE

g3 connectivity: TRUE

g4 connectivity: TRUE

g5 connectivity: TRUE

So we can find that g1 and g2 are disconnected, g3, g4, g5 are connected.

g1 Connectivity: 0

g2 Connectivity: 0

g3 Connectivity: 0.980198

g4 Connectivity: 0.990099

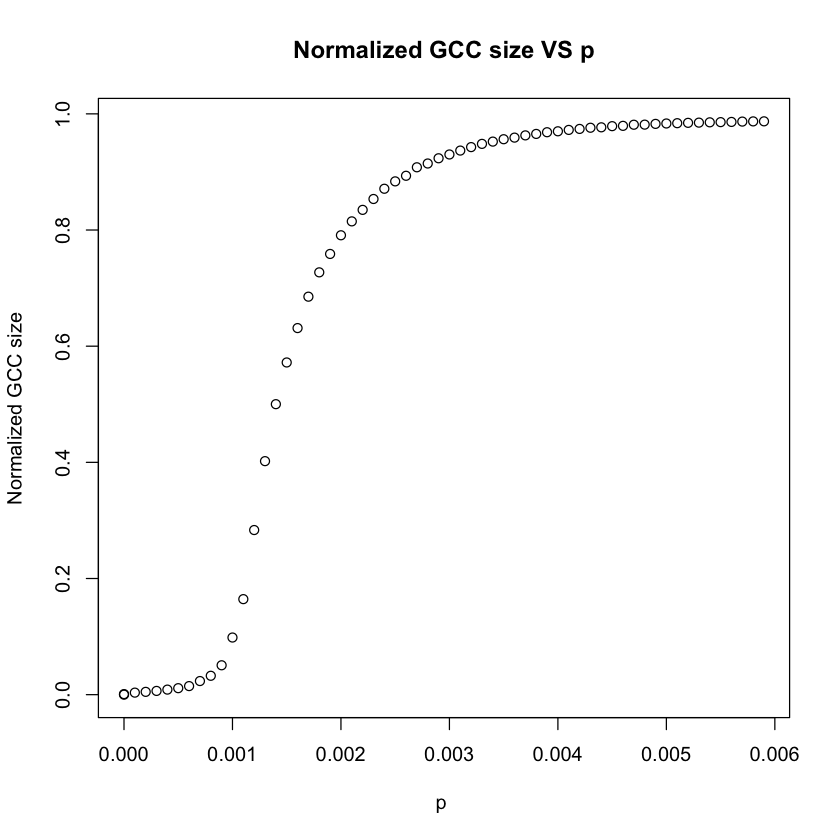
g5 Connectivity: 0.990099

So we can conclude that the Numerical connectivity of g1 and g2 is 0, g3 is 0.98198, g4 and g5is 0.99099. So we need to find GCC of g1 and g2.

For g1: GCC is 935, and diameter is 13

For g2: GCC is 982, and diameter is 12

(c) It turns out that the normalized GCC size (i.e., the size of the GCC as a fraction of the total network size) is a highly nonlinear function of p, with interesting properties occurring for values where p = O( ln n/n ). For n = 1000, sweep over values of p in this region and create 100 random networks for each p. Then scatter plot the normalized GCC sizes vs p. Empirically estimate the value of p where a giant connected component starts to emerge (define your criterion of “emergence”)? Do they match with theoretical values mentioned or derived in lectures?

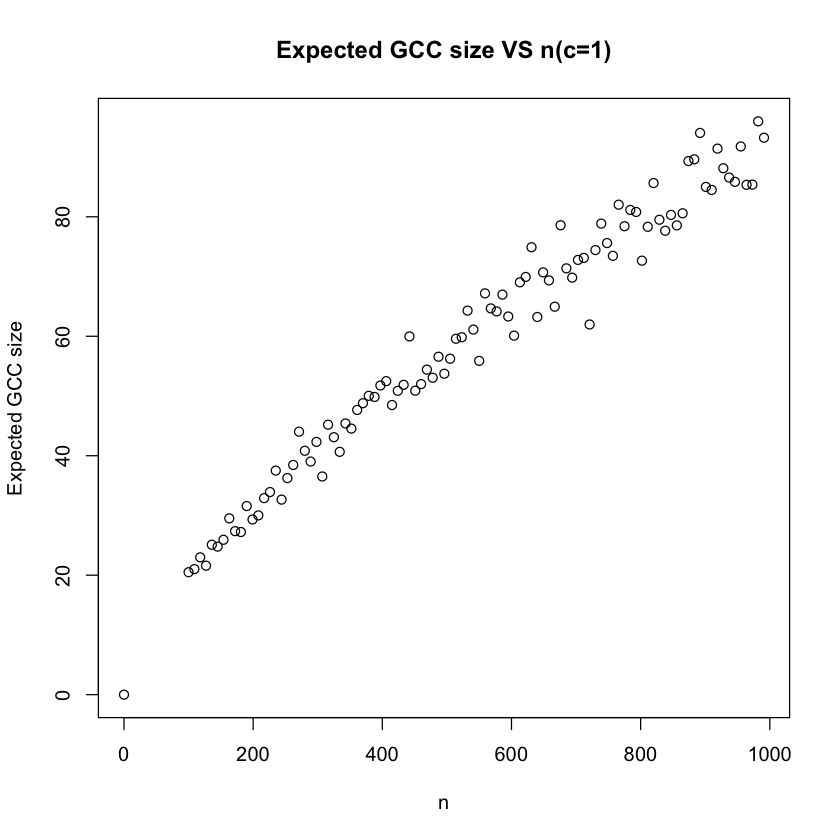
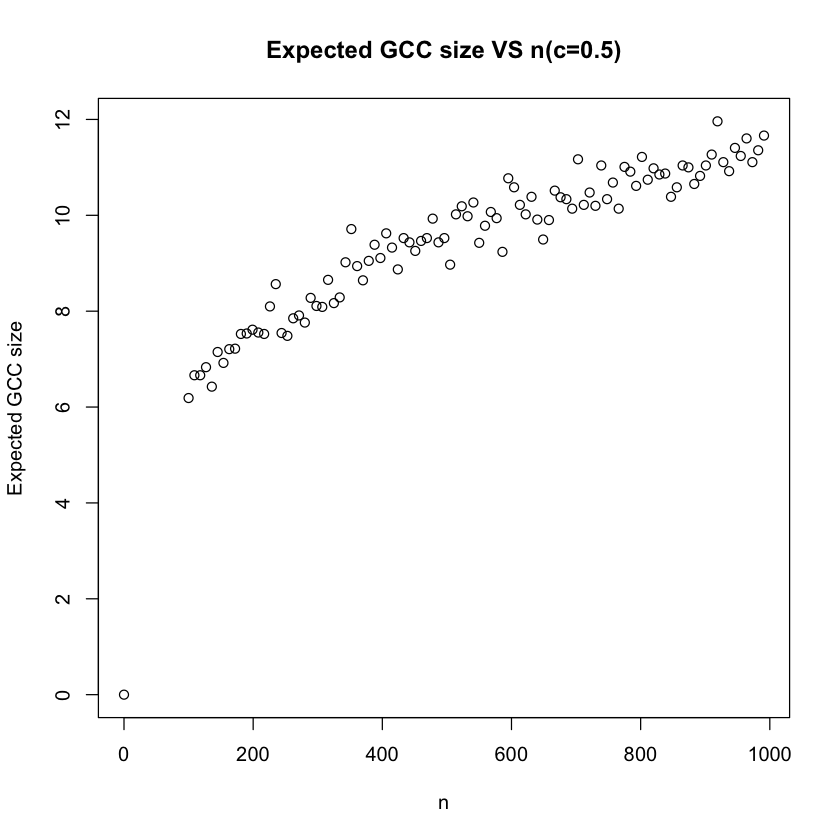


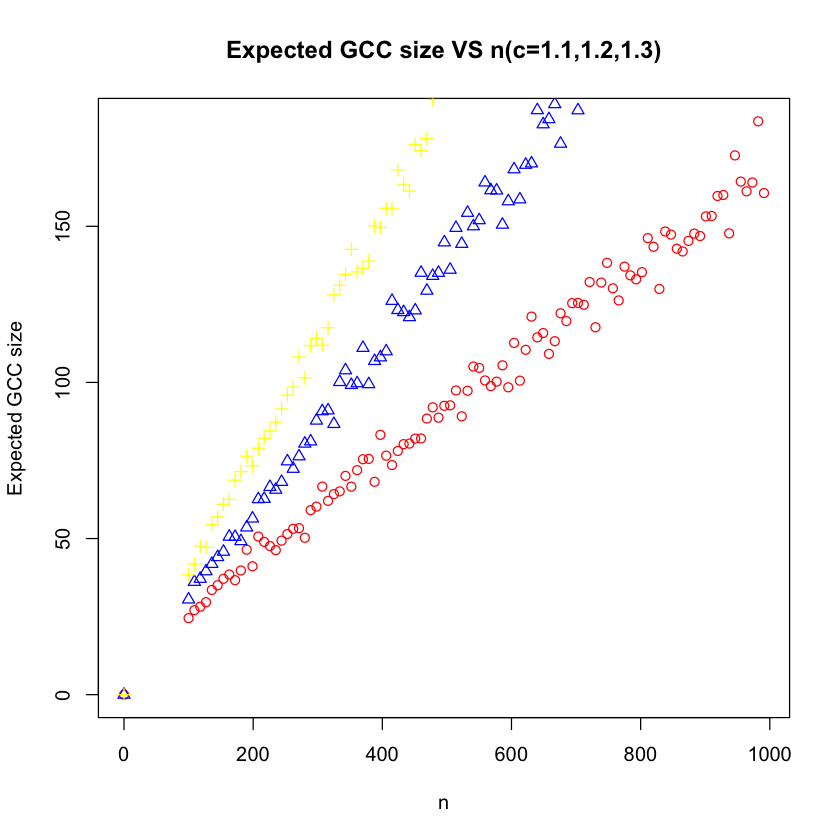
Value of p where a giant connected component starts to emerge is 0.005473913 in the experiment. The theoretical value of p should be ln n/n = 0.0069078. Thus they are basically equal, but the practical value is smaller than the theoretical value.

(d) i. Define the average degree of nodes c = n × p = 0.5. Sweep over number of nodes, n, ranging from 100 to 10000. Plot the expected size of the GCC of ER networks with n nodes and edge-formation probabilities p = c/n, as a function of n. What trend is observed?

ii. Repeat the same for c = 1.

iii. Repeat the same for values of c = 1.1, 1.2, 1.3, and show the results for these three values in a single plot.





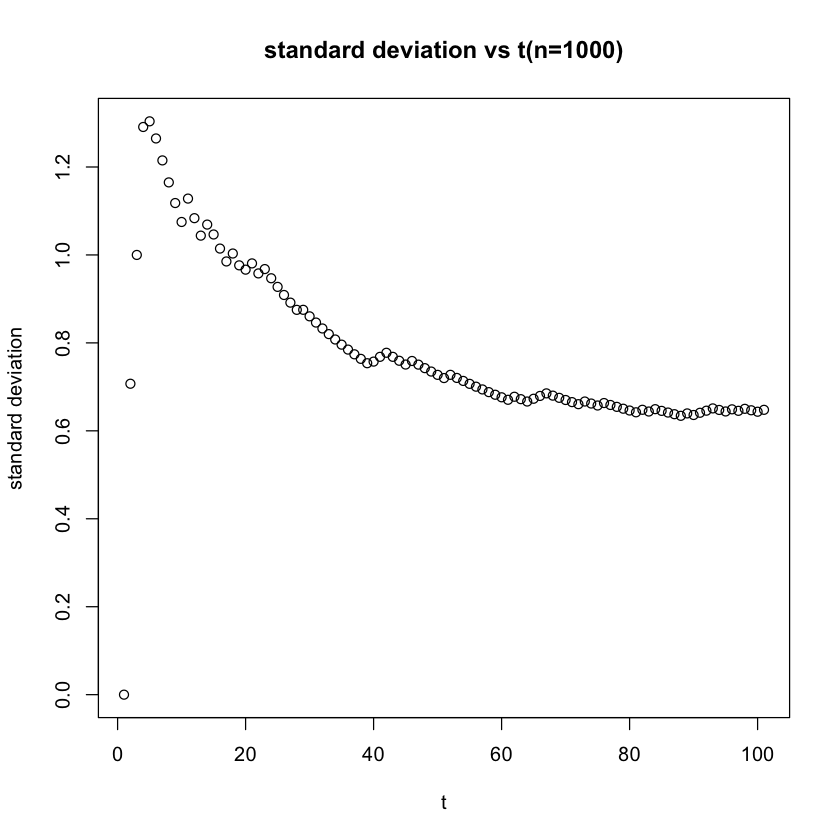
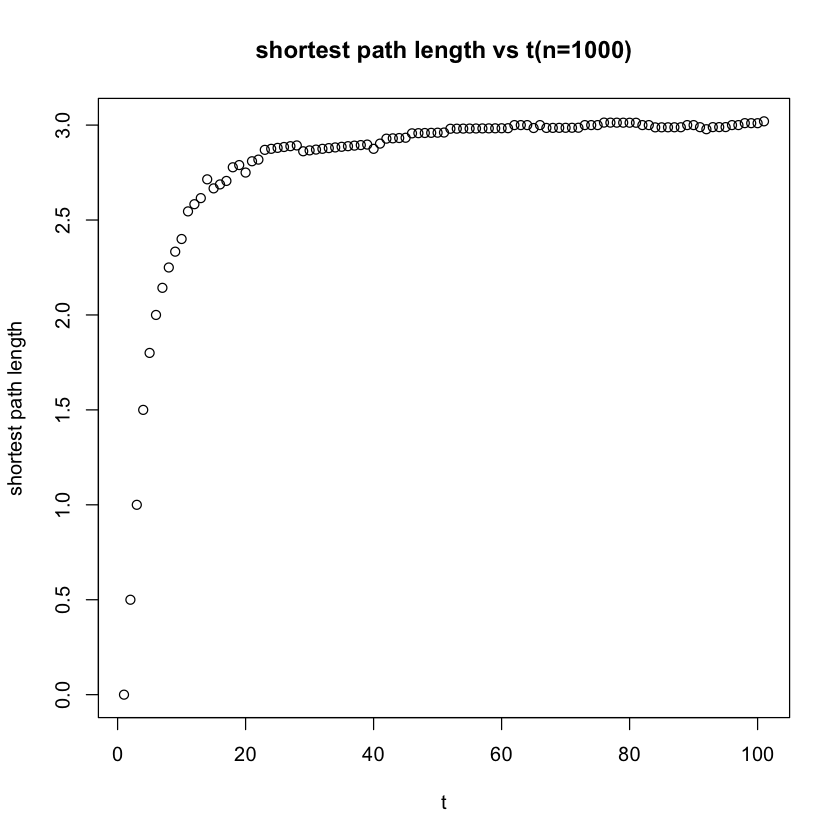
As shown in the plot, expected GCC size is linearly proportional to the number of nodes. In the last figure, the yellow scatter represents c=1.3, blue represents c=1.2, red represents c=1.1. So larger c will lead to larger expected GCC size.

**2 Random Walk on Networks**

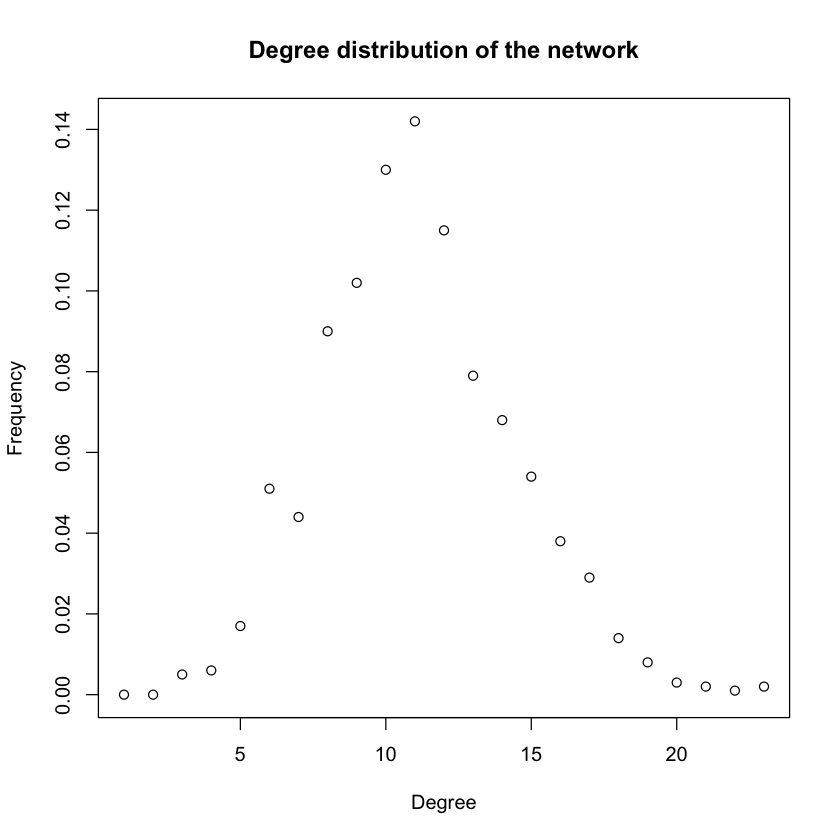
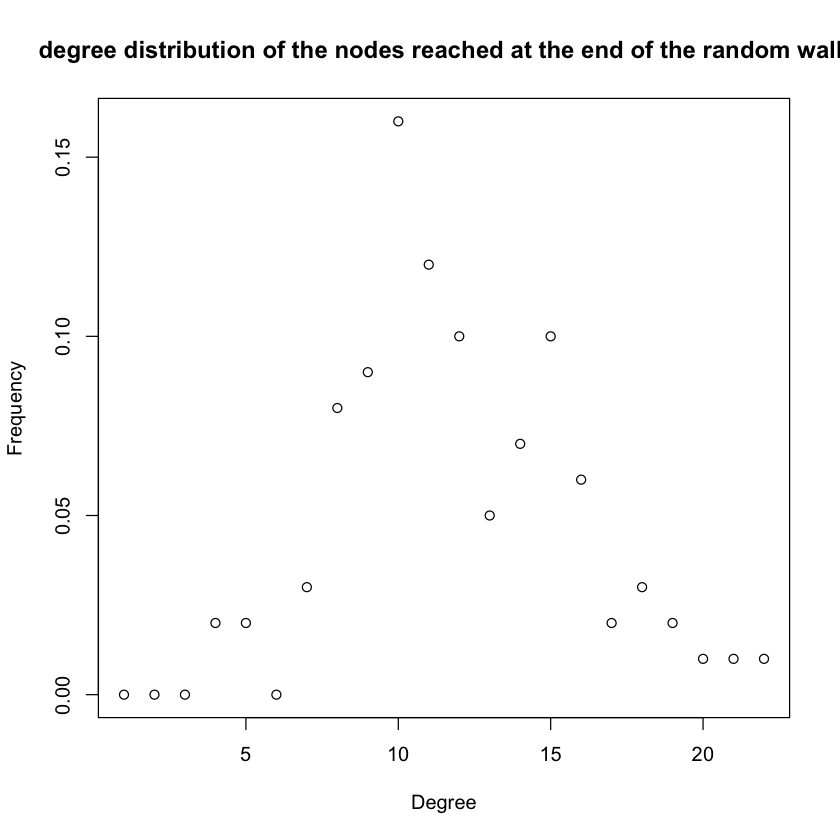
1. Random walk on Erdös-Rényi networks

(a) Create an undirected random network with 1000 nodes, and the probability p for drawing an edge between any pair of nodes equal to 0.01.

(b) Let a random walker start from a randomly selected node (no teleportation). We use t to denote the number of steps that the walker has taken. Measure the average distance (defined as the shortest path length) <(s(t)> of the walker from his starting point at step t. Also, measure the standard deviation σ2 (t) = <(s(t)-<s(t)>)2> of this distance. Plot <(s(t)> v.s. t and σ2(t) v.s. t. Here, the average <·> is over random choices of the starting nodes.

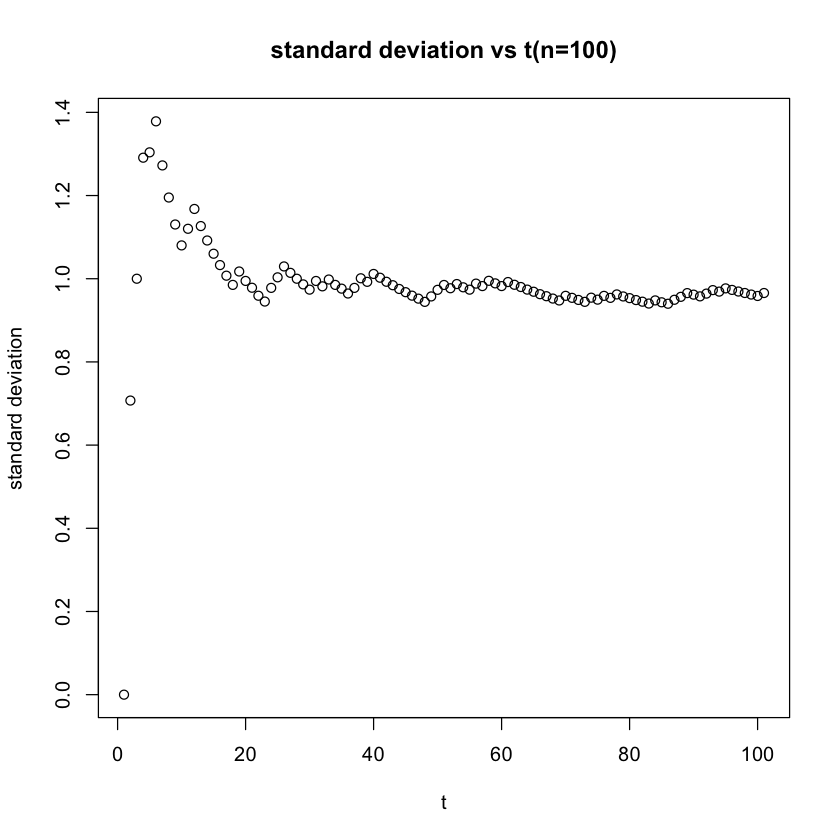
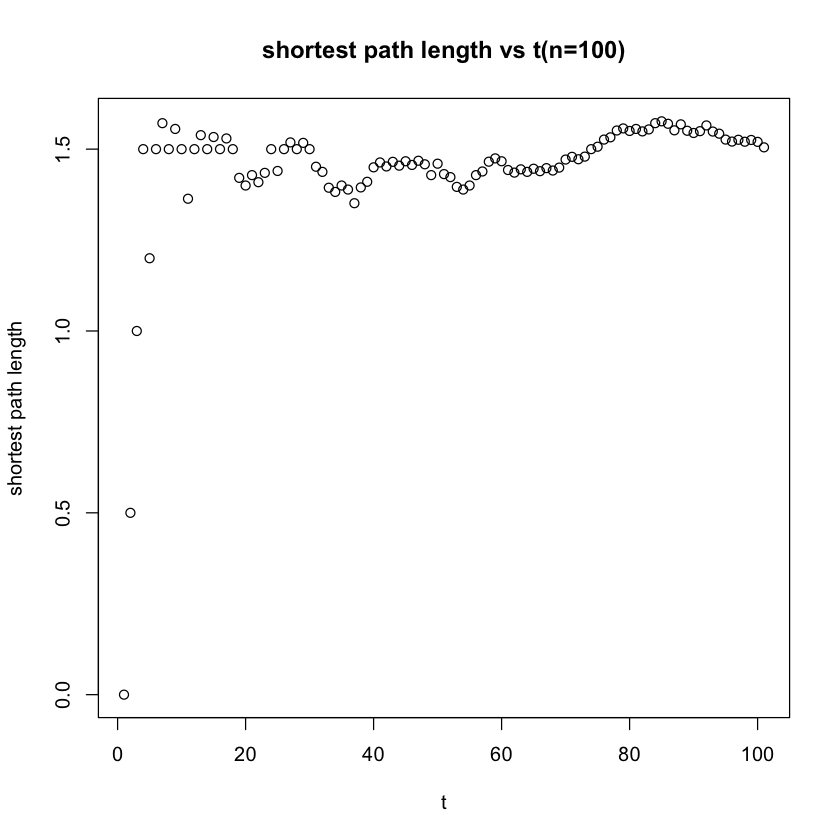


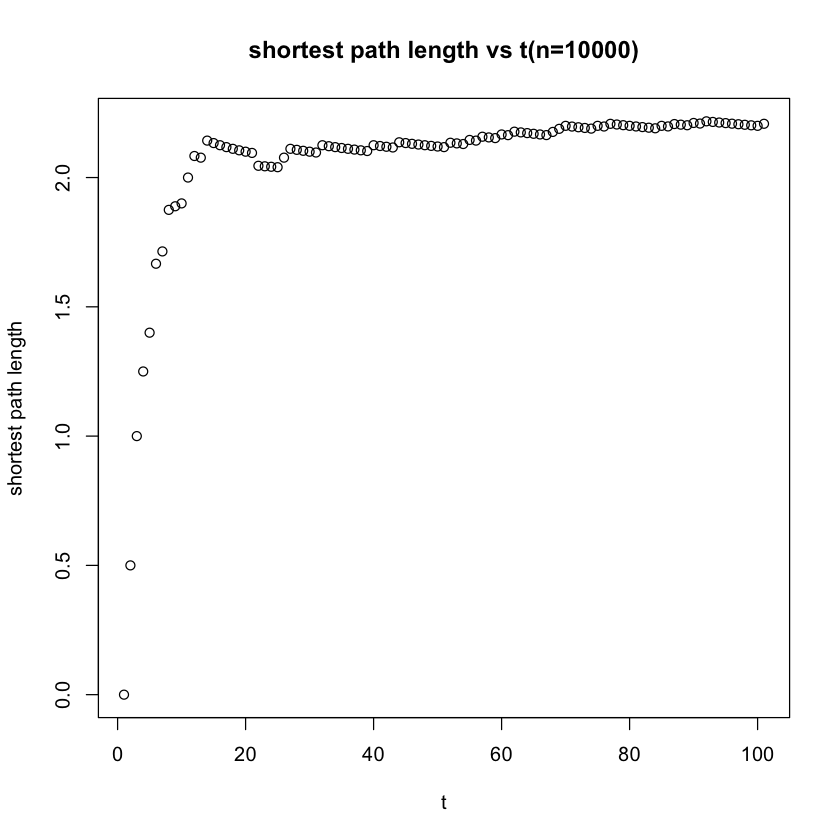
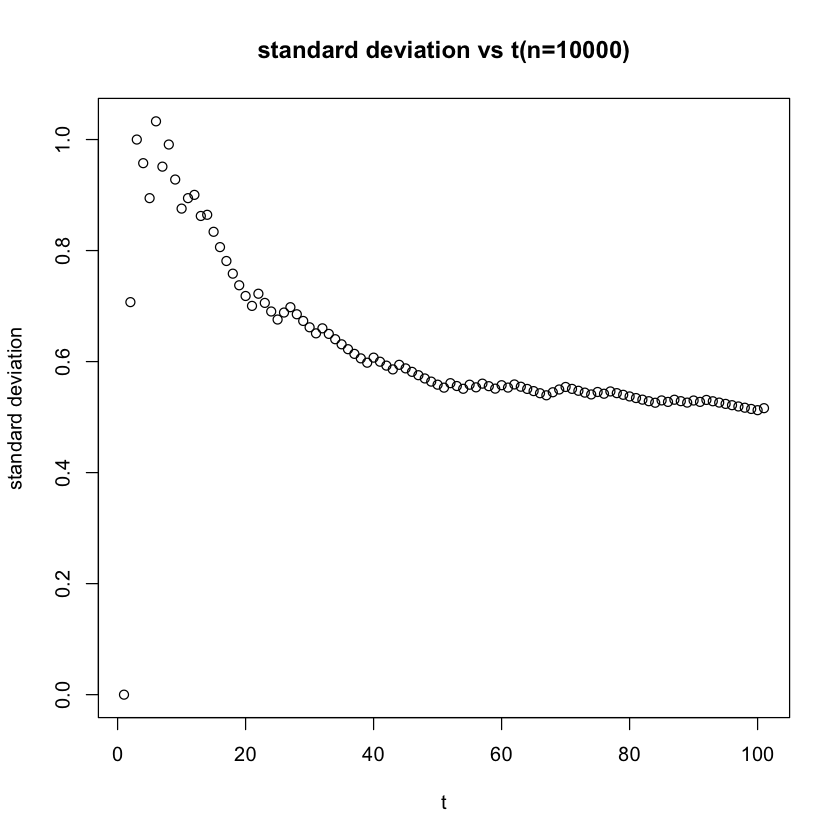
(c) Measure the degree distribution of the nodes reached at the end of the random walk. How does it compare to the degree distribution of graph?



The degree distribution of the nodes reached at the end of the random walk is similar to the degree distribution of graph.

(d) Repeat (b) for undirected random networks with 100 and 10000 nodes. Compare the results and explain qualitatively. Does the diameter of the network play a role?



|  |  |  |  |
| --- | --- | --- | --- |
| nodes | 100 | 1000 | 1000 |
| diameter | 6 | 5 | 3 |