1. C) cluster 1: size=113, common=1,percentage=0.53

cluster 2: size=105, common=3,percentage=0.31

cluster 3: size=149, common=7,percentage=0.46

cluster 4: size=147, common=3,percentage=0.37

cluster 5: size=57, common=2,percentage=0.91

cluster 6: size=44, common=2,percentage=0.64

cluster 7: size=91, common=6,percentage=0.89

cluster 8: size=99, common=1,percentage=0.40

cluster 9: size=135, common=4,percentage=0.44

correct 531 out of 1000. error of 0.47

d) cluster 0: size=291, common=1,percentage=0.10

cluster 1: size=1, common=0,percentage=1.00

cluster 2: size=1, common=2,percentage=1.00

cluster 3: size=1, common=4,percentage=1.00

cluster 4: size=1, common=4,percentage=1.00

cluster 5: size=1, common=5,percentage=1.00

cluster 6: size=1, common=5,percentage=1.00

cluster 7: size=1, common=5,percentage=1.00

cluster 8: size=1, common=6,percentage=1.00

cluster 9: size=1, common=6,percentage=1.00

correct 39 out of 1000. error of 0.87

e) k\_means:

cluster 0: size=228, common=3,percentage=0.38

cluster 1: size=96, common=0,percentage=0.89

cluster 2: size=223, common=1,percentage=0.43

cluster 3: size=114, common=6,percentage=0.70

cluster 4: size=270, common=4,percentage=0.32

cluster 5: size=69, common=2,percentage=0.87

correct 493 out of 1000. error of 0.51

singlelinkage:

cluster 0: size=295, common=0,percentage=0.10

cluster 1: size=1, common=2,percentage=1.00

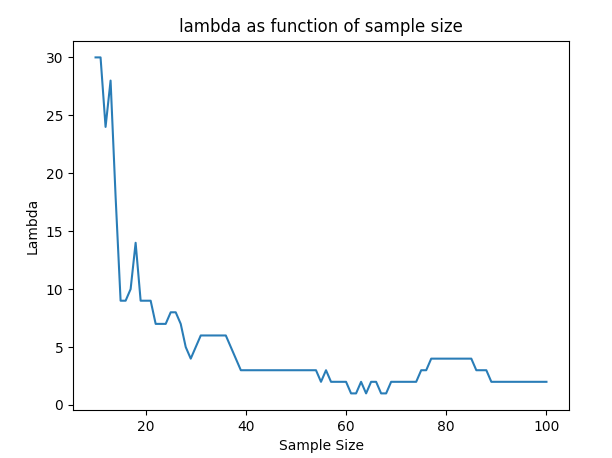
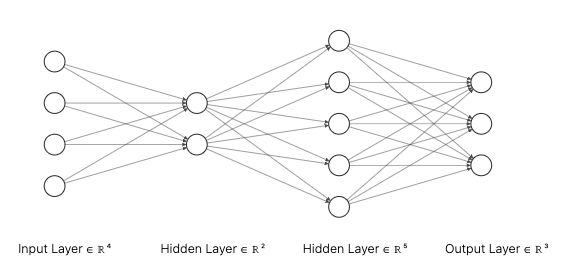
cluster 2: size=1, common=2,percentage=1.00

cluster 3: size=1, common=2,percentage=1.00

cluster 4: size=1, common=3,percentage=1.00

cluster 5: size=1, common=4,percentage=1.00

correct 35 out of 1000. error of 0.88

1. A)   
     
   **B)** We expect to see the value of optimal decreases as the sample size increases. This is because when the sample size is low, the sample does not represent the distribution properly, hence we will obtain large hypothesis class which will result in overfitting. To handle that, a higher will be required as a penalty to reduce the hypothesis class size.   
   As the sample size increases, we expect to see a decrease in the optimal value of , until convergence. As the sample size increases, the hypothesis class size decreases. This results in less overfitting which means we don’t want to penalize the norm of w as much.  
     
   **C)** Yes, this is we got what we expected as explained 2.b, in the plot submitted in 2.a.  
     
   D)
2. A
3. A)    
     
   B) C)   
   D)