Network science

Complex System :- Behind the complex system there is a "network", that defines the interactions between the components

Example :-

1. Facebook
2. Structure of any organization.
3. ~~Neutrons~~ Neurons of Human Brain
4. Business
5. Finance.
6. Internet
7. Human Genes

A Complex System is made up of many non-identical "elements" connected by the "interactions".

Role of Network

We will Never understand complex system unless we map out and understand the network behind them.

Network used in :- The architecture of network emerging in various domains of science, nature & technology.

Networks & Graphs

Components :- nodes, vertices
interactions :- links, edges.
System :- Network, Graph.

Example :- (i) www
(ii) Social Network
(iii) Metabolic Network.

} Network

Language (Network, ~~link~~ node, link)

Graph :- Mathemetical representation of network :-
(i) web graph
(ii) Social graph

Language (~~Graph~~ Graph, vertex, edge)

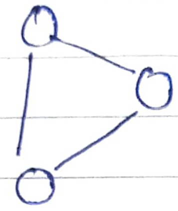
Degree Distribution :-

Average Degree :-

$$\langle K \rangle = \frac{1}{N} \sum_{i=1}^{N} K_i = \frac{2L}{N} \quad (\text{for undirected})$$

where : K_i = ith link.

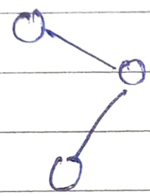
Example



average Degree :-

$$\langle k \rangle = \frac{1}{3}(2+2+2) = 3$$

undirected



$$= \frac{1}{3}\left(\cancel{2+2}\right)^{(1+2+1)} = \frac{4}{3}$$

for Directed :-

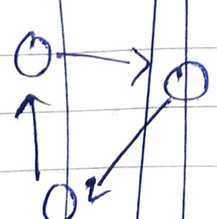
$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k^{\text{in degree}} + \frac{1}{N}\sum_{i=1}^{N} k^{\text{out degree}}$$

Because $\langle k^{in} \rangle = \langle k^{out} \rangle$

the

$$\boxed{\langle k \rangle = \frac{L}{N}}$$

Example :-



$$\langle k^{in} \rangle = \frac{1+1+1}{3} = 1$$

$$\langle k^{out} \rangle = \frac{1+1+1}{3} = 1$$

~~total $\langle k \rangle$~~ = average degree = 2

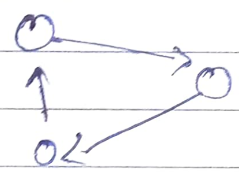
For Directed graph.

Average Digree = ~~Avg~~ $\langle k^{in} \rangle$ or $\langle k^{out} \rangle$

Because: $\langle k^{in} \rangle = \langle k^{out} \rangle$.

$$\boxed{\langle k \rangle = \frac{L}{N}}$$

Example



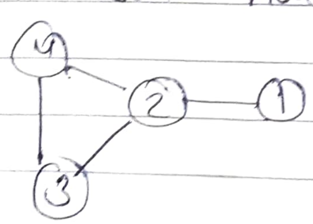
$\langle k^{in} \rangle = \frac{3}{3}$

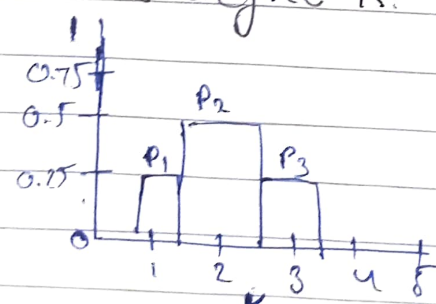
$\langle k^{out} \rangle = \frac{3}{3}$

$$\boxed{\langle k \rangle = 1}$$

Digree Distribution:-

P(K) :- Probablity that a randomly chosen node has degree K.





$P(K) = \frac{N_K}{N}$ → No. of Node with digree K

$P(1) = \frac{1}{4} = 0.25$

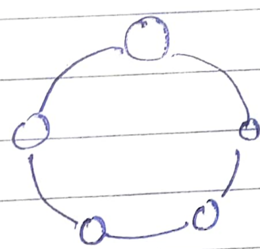
$P(2) = \frac{2}{4} = 0.5$

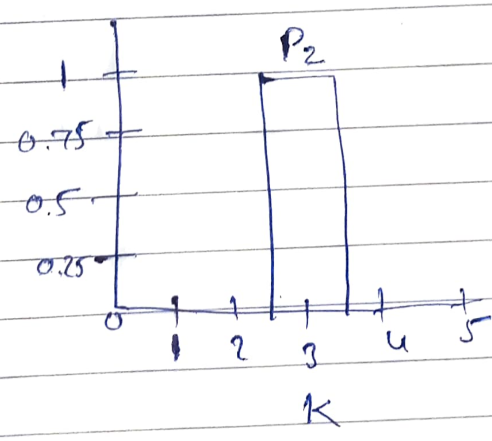
$P(3) = \frac{1}{4} = 0.25$

$P(4) = 0$

$P(5) = 0$

Example 2





$P(1) = 0$

$P(2) = \frac{5}{5} = 1$

$P(3) = 0$

$P(4) = 0$

$P(5) = 0$

Discrete Representation :-

P_k :- The Probablity that the node Have degree 'k'

$$\boxed{P_k = \frac{N_k}{N}}$$

Continous Represantation :-

$$\boxed{\cancel{P(K)} = \int_{K_1}^{K_2} p(K)\, dK} \quad \text{——— (1)}$$

P(K) is the pdf of the digree

& the equation represents the Probablity that a node's degree is between K_1 and K_2

Normalized Condition :-

$$\boxed{\sum_{0}^{\infty} P_K = 1} \quad \text{or} \quad \boxed{\int_{K_{min}}^{\infty} P(K)\, dK = 1}$$

Where K_{min} is the minimal degree in the network.

In our case it is '0'

& K can varis to ∞

Undirected Network Example:-

1. Actor network
2. Protein interactions.

Directed Network Example:-

1. URLs on the WWW
2. Phone calls
3. Metabolic reactions

Representation :-

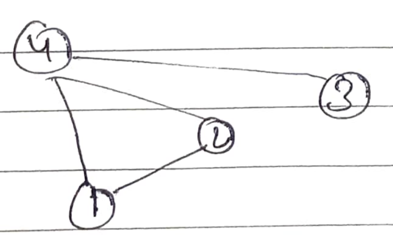
Adjacency Matrix.

$A_{ij} = 1$ if there is a link between i & j

$A_{ij} = 0$ if there is No link between i & j

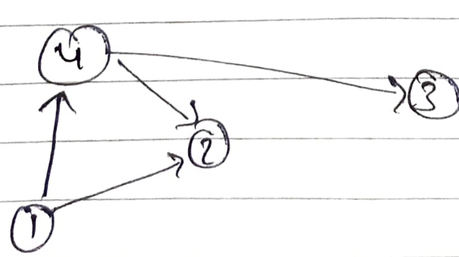
Example :-

Undirected :-



$$A_{ij} = \begin{array}{c c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

Directed :-



$$A_{ij} = \begin{array}{c c} & \begin{matrix} 1 & 2 & 3 & 4 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

No. of links (L) =

$$L = \frac{1}{2}\sum_{i,j}^{N} A_{ij} \quad \text{(Undirected graph)}$$

$$L = \sum_{ij}^{N} A_{ij} \quad \text{(Directed graph)}$$

Max Links in a Network :-

$$\boxed{L_{max} = {}^{N}C_{2} = \frac{(N)(N-1)}{2}} \quad , \quad \left\{ \langle K \rangle = \frac{2L}{N} \right.$$

A graph with degree $L = L_{max}$ is a complete graph.

& Its average degree $\boxed{\langle K \rangle = N-1}$

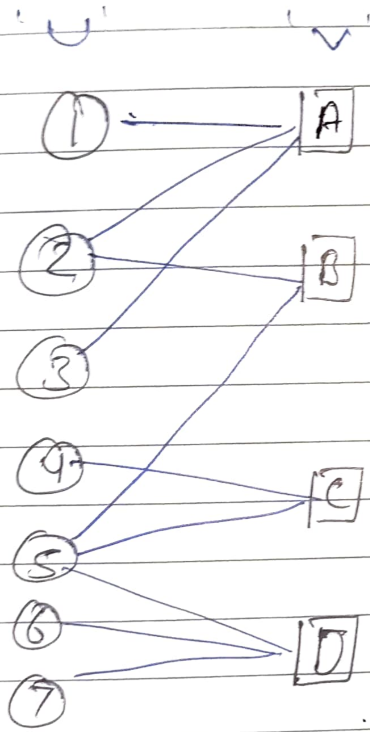
Most of the network has

$$\boxed{L \ll L_{max}}$$

&

$$\boxed{\langle K \rangle \ll (n-1)}.$$

Bipartite graph:- It is a graph whose nodes can be devided into two disjoint sets 'U' & 'V' Such that every link in a graph connects nodes in the set 'U' to the nodes in the set 'V'.



① Disease Network
② Collaboration Network.

A Path is a sequence of nodes in which each node is adjacent to next one.

Path

~~Path~~ :- The distance between two nodes is defined as the number of edges along the sortest path connecting them.

In directed graphs each path needs to follow the direction of arrows.
Thus in a digraph the distance from node A to B is generally different from the node B to A

Number of Paths between any two nodes i & j

Denoted By $N_{ij}^{(L)}$

$\downarrow$

Number of Path of length L from i to j.

$N_{ij}^{(1)}$ = If there is a link between i & j then $\boxed{A_{ij} = 1}$ else $\boxed{A_{ij} = 0}$

$N_{ij}^{(2)}$ = If there is a path of length two between i & j then

$\boxed{A_{ik} A_{kj} = 1}$ else $\boxed{A_{ik} A_{kj} = 0}$

Then the Number of path of length 2

$$N_{ij}^{(2)} = \sum_{k=1}^{N} A_{ik} A_{kj} = [A^2]_{ij}$$

$N^{(n)}$:- The Number of path of length 'n' from i to j

If there is a Path of length 'n' b/w ~~from~~ i & j then $A_{ik} \ldots A_{ij} = 1$

otherwise

$$\boxed{A_{ik} \cdots A_{lj} = 0}$$

then $\boxed{N_{ij}^{(n)} = [A^n]_{ij}}$

Find Distance :- (use BFS Algorithm)

Network Diameter And Average Distance:-

Diameter (d_{max}) = The maximum distance between any pair of nodes in the graph.

Average Path length / distance $\langle d \rangle$ for connected graph:-

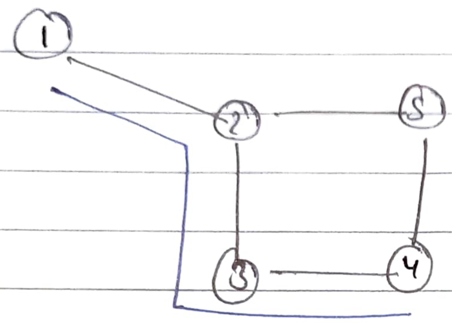
$$\boxed{\langle d \rangle = \frac{1}{2L_{max}} \sum_{i,j \neq i} d_{ij}}$$

In an undirected graph $d_{ij} = d_{ji}$, so we only need to count them once.

$$\langle d \rangle = \frac{1}{} \sum_{i,j>i} d_{ij}$$

Example :-

Diameter :-



The longest Path in a graph

Average Path length :-

$$\frac{\text{Sum of length of } \cancel{\text{each}} \text{ path b/w each pair of node}}{\text{total no. of Path, b/w each Pair of Nodes}}$$

$$\frac{[d_{1\to2} + d_{1\to3} + d_{1\to4} + d_{1\to5} + d_{2\to3} + d_{2\to4} + d_{2\to5} + d_{3\to4} + d_{3\to5} + d_{4\to5}]}{10}$$

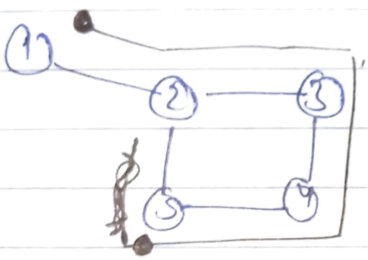
$$\Rightarrow \frac{1+2+3+2+1+2+1+1+2+1}{10}$$

$$\Rightarrow \frac{16}{10} = 1.6$$

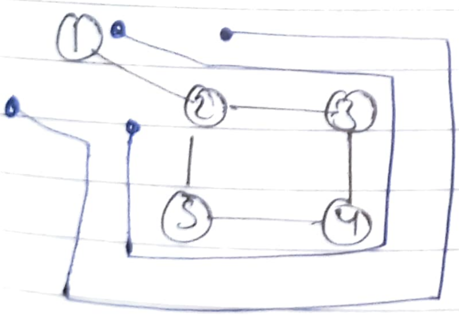
Cyclic Path :- A path with same start & end node.

2, 3, 4, 5 cycle

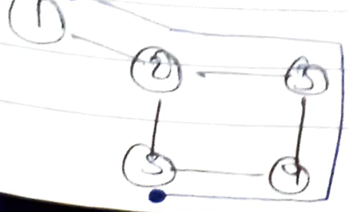
Self-avoiding path :- A path that does not intersect itself.

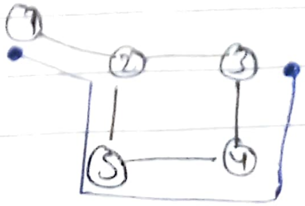


Eulerian path :- A Path that traverse each link exactly once.

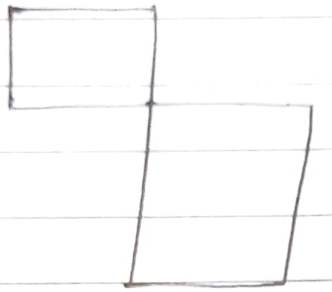


Hamiltonian Path :- A path that visits each node exactly once





If the adjcency list is like



then there is a connected componant

Strongly connected Component:-

A graph has a node from each node to every other node & vice versa.

Clustering Coefficient

This clustering coefficient gives what fraction of your neighbors are connected.

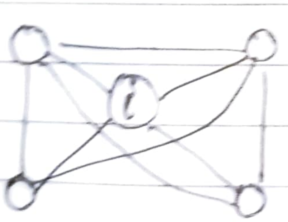
Denoted by

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

$C_i \in \{0,1\}$

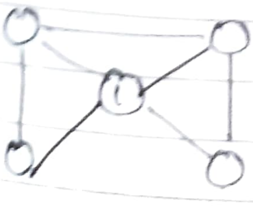
e_i = no of edges between the neighbour of node i

Example :-



k_i = degree of node i

$$C_i = \frac{2\times 6}{4\times 3} = 1$$



$$C_i = \frac{2\times 3}{4\times 3} = \frac{1}{2}$$

Clustering Coefficient

This clustering cofficient gives what fraction of your neighbors are connected.

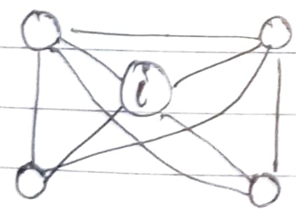
Denoted by

$$\boxed{C_i = \frac{2e_i}{K_i(K_i-1)}}$$

$C_i \in \{0,1\}$

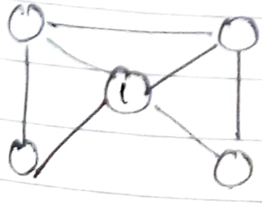
e_i = no of edges between the neighbour of node i

Example :-

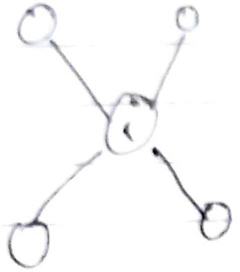


K_i = digree of node i

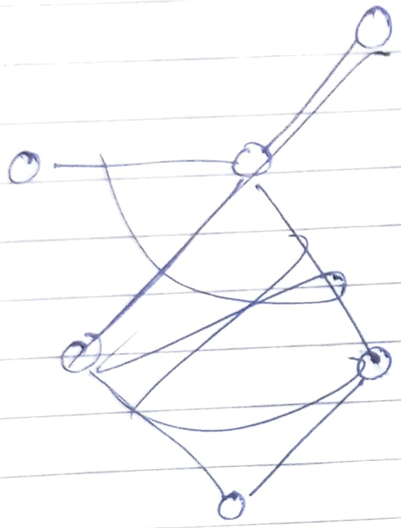
$$C_i = \frac{2 \times 6}{4 \times 3} = 1$$

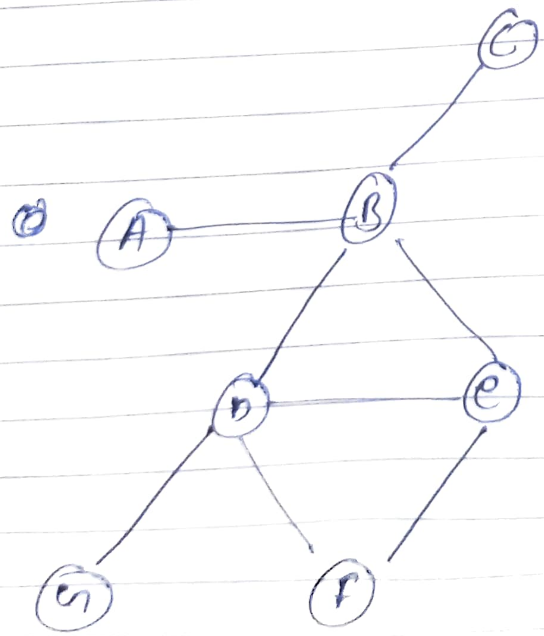


$$C_i = \frac{2 \times 3}{4 \times 3} = \frac{1}{2}$$

1 

$$C_i = \frac{2\times 0}{4\times 3} = 0$$





Node	C_i
A	$\frac{2\times 0}{1\times 0} = 0$
B	$\frac{2\times 1}{4\times 3} = \frac{1}{6}$
c	$\frac{2\times 0}{1\times 0} = 0$
D	$\frac{2\times 2}{4\times 3} = \frac{1}{3}$
e	$\frac{2\times 2}{3\times 2} = \frac{2}{3}$
G	$\frac{2\times 0}{1\times 0} = 0$
f	$\frac{2\times 1}{2\times 1} = 1$

Random Network model

Definition :- A random graph is a graph of N nodes where each pair of nodes is connected by probablity p.

$P(L)$:- The Probablity to have exactly L Links in a network of N nodes and probability p :

Binomial Distribution

$$\boxed{P(L) = \binom{L_{max}}{L} p^L (1-p)^{L_{max}-L}}$$

$$L_{max} = \frac{n(n-1)}{2}$$

$$\boxed{P(L) = \binom{\frac{n(n-1)}{2}}{L} p^L (1-p)^{\frac{n(n-1)}{2}-L}}$$

No. of different ways we can choose L links among all the links

Binomial Distribution:-

$P(n) = {}^{N}C_{n}\, p^{n} (1-p)^{N-n}$ (Probability p)

first Moment mean $\langle n \rangle = NP$

Second Moment $\langle n^2 \rangle = P(1-P)N + P^2N^2$

Variance $= \underbrace{E[n^2]}_{\text{Second moment}} - \underbrace{(E[n])^2}_{\text{first Moment}}$

$= P(1-P)N + \cancel{P^2}N^2 - \cancel{NP}^2$

$= \cancel{PN - P^2N + P^2N^2 - NP}$

$$\boxed{\sigma_n^2 = P(1-P)N} \Rightarrow NPq$$

$\boxed{\sigma_n = [P(1-P)N]^{1/2}} \Rightarrow \sqrt{NPq}$

$P(L)$: Probablity that the network has L links

$$\boxed{P(L) = {}^{\frac{n(n-1)}{2}}C_L \; p^L (1-P)^{\frac{n(n-1)}{2} - L}}$$

Average number of links $\langle L \rangle$ in a random network is given by

$$\boxed{\langle L \rangle = \sum_{L=0}^{\frac{n(n-1)}{2}} L \cdot P(L) = P \cdot \frac{N(N-1)}{2}}$$

Average digree $\langle K \rangle$ in random graph

$$\langle K \rangle = \frac{2L}{N} = \frac{\cancel{2}}{N} \times \frac{P \times N(N-1)}{\cancel{2}}$$

$$\boxed{\langle K \rangle = P(N-1)}$$

σ^2 (variance) $= NPq$

$$= \frac{n(n-1)}{2} \times P(1-P)$$

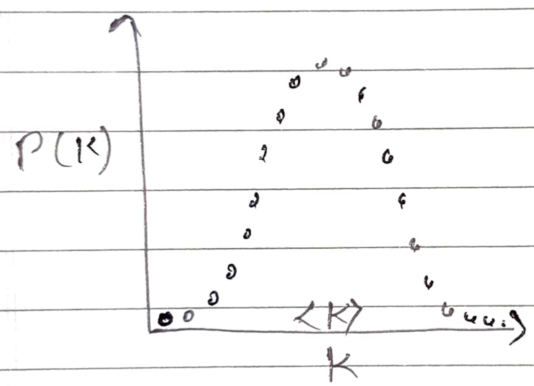
$$\boxed{\sigma^2 = P(1-P) \frac{N(N-1)}{2}}$$

In random graph the average link is

$$P(N-1)$$

But in Normal graph it is $(N-1)$

Digree Distribution



$P(K)$ = Probablity ~~the~~ ~~of~~ that the random graph has Nodes with digree K

$$\boxed{P(K) = {}^{N-1}C_K \; p^K \; (1-p)^{(N-1)-K}}$$

Selecting K nodes from N-1

Probablity of Having K edges

Probablity of missing N-1-K edges

$$\langle K \rangle = P(N-1) \quad \& \quad \sigma_k^2 = P(1-P)(N-1)$$

$$\frac{\sigma_K}{\langle K \rangle} = \sqrt{\frac{\sigma_K^2}{\langle K \rangle^2}} = \left[\frac{\cancel{P}(1-P)\cancel{(N-1)}}{P^2(N-1)^2}\right]^{1/2}$$

$$\frac{\sigma_K}{\langle K \rangle} = \left[\frac{(1-P)}{P} \times \frac{1}{N-1}\right]^{1/2} \approx \frac{1}{(N-1)^{1/2}}$$

for $n \longrightarrow \infty$

$\quad K \longrightarrow 0$

then the Binomial Distribution did Not work

So we have to convert the Binomial Distribution to poision distribution.

$$P(K) = {}^{N-1}C_K \, P^K (1-P)^{N-1-K}$$

Solving each part saperatly

$${}^{N-1}C_K = \frac{(N-1)!}{K!\,(N-1-K)!}$$

$$= \frac{(N-1-1)(N-1-2)(N-1-3)\ldots(N-1-(K+1))\cancel{(N-1-K)!}}{\lfloor K \; \cancel{\lfloor N-1-K}}$$

$$= \frac{(N-1-1)(N-1-2)(N-1-3)\ldots\ldots(N-1-(K+1))}{\lfloor K}$$

$$= \frac{(N-1)^K}{\lfloor K}\left[\underbrace{\left(1-\frac{1}{N-1}\right)\left(1-\frac{2}{N-1}\right)\ldots\ldots\left(1-\frac{K+1}{N-1}\right)}_{\approx 1}\right]$$

$\because \quad N \to \infty$

$$\boxed{\binom{N-1}{K} = \frac{(N-1)^K}{\lfloor K}}$$

Now solving :-

$= (1-P)^{N-1-K}$

~~take~~ ~~by~~ ~~we~~

we can write above expression as

$$= e^{\log_e (1-P)^{N-1-K}}$$

$$= e^{N-1-K \log(1-P)}$$

$$= \cancel{e^{N-1-K}}$$

$$\left\{ \because \log(1-P) = -\left[P + \frac{P}{2} + \frac{P}{3} \ldots \infty\right] \right.$$

$$= e^{-(N-1-K)\left[P + \frac{P^2}{2} + \frac{P^3}{3} \ldots \infty\right]}$$

$$= e^{-(N-1-K)\, P \underbrace{\left[1 + \frac{P}{2} + \frac{P^2}{3} \ldots \infty\right]}_{\approx 1}}$$

$$\left\{ \because P \to 0 \right.$$

$$\boxed{(1-P)^{N-1-K} \cong e^{-(N-1-K)P}}$$

$$\left\{ \begin{array}{l} \because \langle K \rangle = P(N-1) \\ P = \dfrac{\langle K \rangle}{(N-1)}. \end{array} \right.$$

$$(1-p)^{N-1-K} = e^{-(N-1-K)\frac{\langle K\rangle}{N-1}}$$

$$= e^{-\langle K\rangle\left(1-\frac{K}{N-1}\right)}$$

$\because N \to \infty$

$$= e^{-\langle K\rangle(1-0)}$$

$$\boxed{(1-p)^{N-1-K} = e^{-\langle K\rangle}}$$

Put all things in our actual equation.

$$\boxed{{}^{n-1}C_K \, p^K (1-p)^{N-1-K} = \frac{(N-1)^K}{\lfloor K} \times e^{-\langle K\rangle}}$$

Binomial to poission.

for $N \to \infty$

$K \to 0$.

$\therefore (N-1) = \frac{\langle K\rangle}{p}$

$$\text{or} = e^{-\langle K\rangle} \frac{\langle K\rangle^K}{p^K \lfloor K}$$

for Poission distribution

$$\boxed{P(K) \cong e^{-\langle K \rangle} \frac{\langle K \rangle^{K}}{\lfloor K}}$$

Distance in a random graph:-

$\langle K \rangle^{1}$ = Nodes at distance one ($d=1$)
$\langle K \rangle^{2}$ = Nodes at distance two ($d=2$)
$\langle K \rangle^{3}$ = Nodes at distance three ($d=3$)
⋮
$\langle K \rangle^{d}$ = Nodes at distance d

$$N = 1 + \langle K \rangle + \langle K \rangle^{2} + \langle K \rangle^{3} \cdots \langle K \rangle^{d_{max}}$$

$$N = \frac{\langle K \rangle^{d_{max}+1} - 1}{\langle K \rangle - 1}$$

$$N \cong \langle K \rangle^{d_{max}}$$

$$\log N = d_{max} \cdot \log \langle K \rangle$$

$$\boxed{d_{max} = \frac{\log N}{\log \langle K \rangle}}$$

Or

$$\boxed{\langle d \rangle = \frac{\log N}{\log \langle K \rangle}}$$

Average distance b/w 2 randomly chosen nodes.

$\frac{1}{\log \langle K \rangle}$ implice that denser the network, the smaller the the distance b/w the nodes

Clustering Cofficient for random graph:-

$$\boxed{C_i = \frac{2\langle L \rangle}{K_i(K_i-1)} = p = \frac{\langle K \rangle}{N}}$$

C decreases with the System size N.

Average path length for random graph

$$\boxed{\langle l_{rand} \rangle = \frac{\log N}{\log \langle K \rangle}}$$

Clustering Coefficient for random graph.

$$\boxed{C_i = \frac{2\langle L \rangle}{K(K-1)} = p = \frac{\langle K \rangle}{N}}$$

Degree Distribution For random graph :-

$$P(K) = e^{-\langle K \rangle} \frac{\langle K \rangle^K}{K!}$$