HOMEWORK 1

<u>MSSV: 20127054</u>

<u>Họ tên: Ngô Văn Trung Nguyên</u>

1. Phát sinh mạng Erdős-Rényi

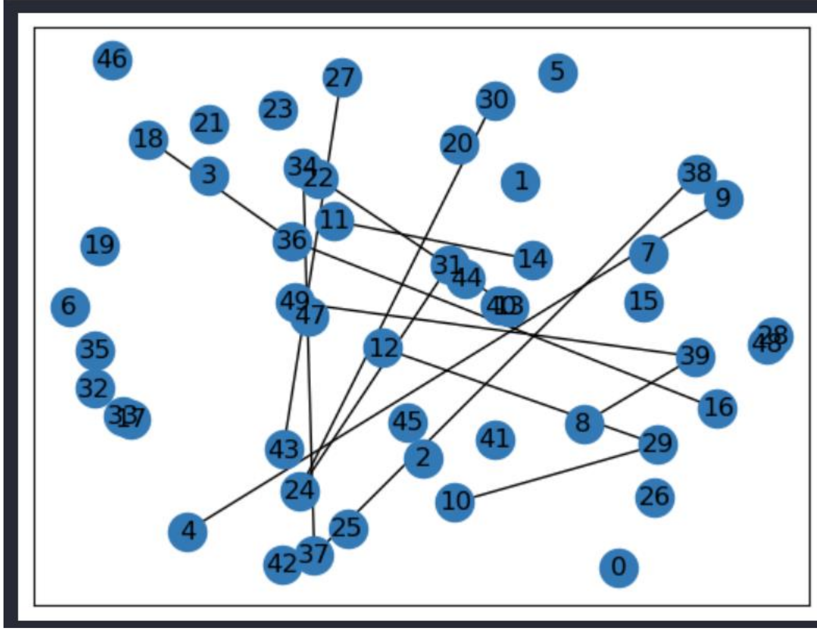
Hãy trực quan hoá mạng Erdős-Rényi với N = 50 nút và bậc trung bình $\langle k \rangle$ lần lượt là

a. $\langle k \rangle = 0.5$

```python
N = 50
k = 0.5
p = k / (N - 1)

G = nx.erdos_renyi_graph(N, p)
nx.draw_networkx(G, nx.random_layout(G, seed=10), with_labels=True)
plt.show()
```

✓ 0.3s

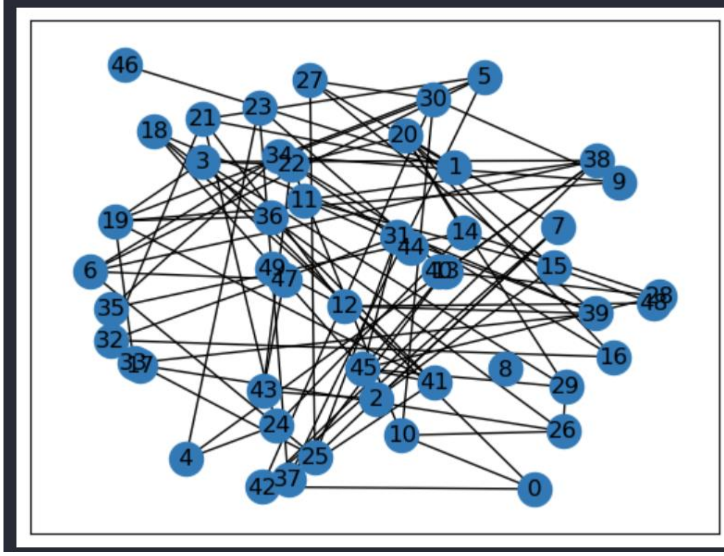


b. $\langle k \rangle = 4$

```python
N = 50
k = 4
p = k / (N - 1)

G = nx.erdos_renyi_graph(N, p)
nx.draw_networkx(G, nx.random_layout(G, seed=10), with_labels=True)
plt.show()
```

✓ 0.2s

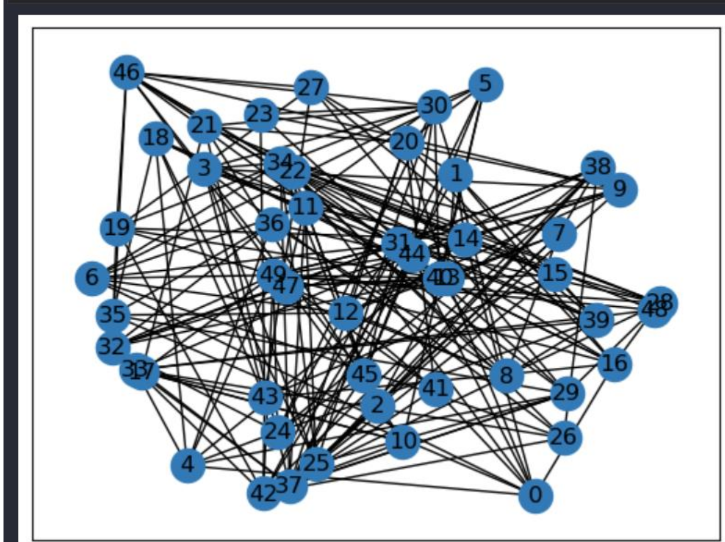


c. $\langle k \rangle = 8$

```python
N = 50
k = 8
p = k / (N - 1)

G = nx.erdos_renyi_graph(N, p)
nx.draw_networkx(G, nx.random_layout(G, seed=10), with_labels=True)
plt.show()
```

✓ 0.2s



2. Mạng Erdős-Rényi

Xem xét mạng G(N, p) có N = 3000 nút và được kết nối với nhau với xác suất p = 10^{-3}

a. *Xác định số lượng liên kết kỳ vọng $\langle L\rangle$ và bậc trung bình $\langle k\rangle$ của mạng*
  - Ta có các công thức sau:
    - Số lượng liên kết kỳ vọng $\langle L\rangle$: $\frac{N(N-1)}{2} * p$
    - Bậc trung bình $\langle k\rangle$: (N-1)*p

  - Thay N = 3000 và p = 10^{-3} vào công thức:
    - $\langle L\rangle = \frac{3000(3000-1)}{2} * 10^{-3} = \mathbf{4498.5}$
    - $\langle k\rangle$ = (3000-1)* 10^{-3} = **2.999**

b. *Xác suất có chính xác 50 liên kết trong mạng là bao nhiêu ?*
  - Ta có công thức sau: $p_L = C^L_{\frac{N(N-1)}{2}} * p^L * (1-p)^{\frac{N(N-1)}{2} - L}$
  - Thay L = 50 vào: $\boldsymbol{p_{50} = C^{50}_{\frac{3000(3000-1)}{2}} * (10^{-3})^{50} * (1-10^{-3})^{\frac{3000(3000-1)}{2} - 50}}$

c. *Dựa vào hình 1, xác định xem mạng ở chế độ (regime) nào ?*
  Do $\langle k\rangle$ = 2.999 nên mạng ở chế độ **Supercritical Regime**.

d. *Tính xác suất p_C để mạng ở chế độ critical point.*
  Để mạng ở chế độ critical point thì $\langle k\rangle = 1$
  Ta có: pc = $\frac{1}{N} = \mathbf{\frac{1}{3000}}$

e. *Tính số nút N^{cr}, bậc trung bình $\langle k^{cr}\rangle$ và khoảng cách trung bình giữa hai nút được chọn ngẫu nhiên $\langle d\rangle$ để mạng chỉ có một thành phần*
  - Tính số nút N^{cr} để mạng chỉ có một thành phần
    - Theo như hình thì mạng ở chế độ Connected Regime thì sẽ có một single giant component và không có isolated nodes hay clusters, nên

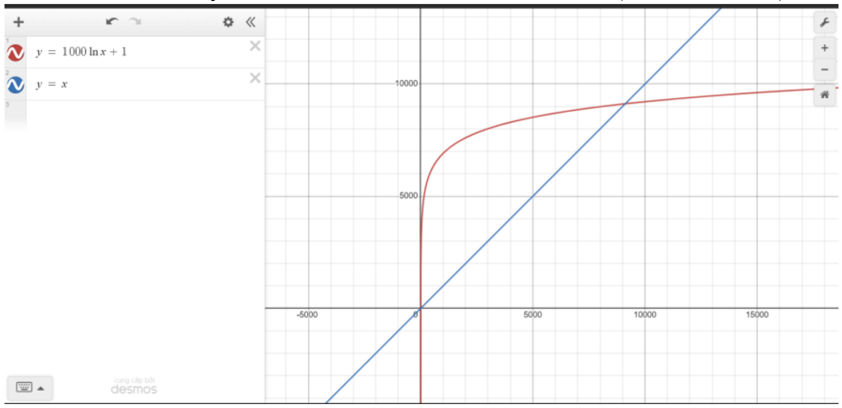
$$\langle k^{cr}\rangle \gg \ln N^{cr}$$

$$\Leftrightarrow (N^{cr}-1) * p \gg \ln N^{cr}$$
$$\Leftrightarrow (N^{cr}-1) * 10^{-3} \gg \ln N^{cr}$$

$$\Leftrightarrow N^{cr} - 1 \gg 1000 \ln N^{cr}$$

$$\Leftrightarrow \ N^{cr} \gg \ 1000 \ln N^{cr} \ + \ 1$$

- Vẽ đồ thị, ta thấy $N^{cr} \gg \ 9120$ hoặc $N^{cr} \ = \ 1$ (do $N^{cr} \geq \ 1$)



- Vậy $\mathbf{N^{cr} \gg \ 9120}$ **hoặc** $\mathbf{N^{cr} \ = \ 1}$ thì mạng chỉ có một thành phần

- Bậc trung bình: $\langle k^{cr} \rangle = (N^{cr} - 1) * p \gg (9120 - 1) * 10^{-3}$

$$\boldsymbol{\langle k^{cr} \rangle \gg \ 9.199}$$

- Khoảng cách trung bình: $\boldsymbol{\langle d \rangle \ = \ \frac{\ln N^{cr}}{\ln \langle k^{cr} \rangle} \ = \ \frac{\ln N^{cr}}{\ln((N^{cr}-1)*p)}}$ **(tùy thuộc vào $\mathbf{N^{cr}}$)**

f. Tìm phân bố bậc p_k của mạng này (xấp xỉ với phân bố bậc Poisson).

- Mạng này là mạng G(N, p) có N = 3000 nút và được kết nối với nhau với xác suất $p \ = 10^{-3}$, đã tính được $\langle k \rangle = 2.999$
  - Binomial distribution: $p_k \ = \ C_{N-1}^{k} \ * \ p^k \ * \ (1-p)^{N-1-k}$
  - Possion distribution: $p_k \ = \ e^{-\langle k \rangle} \ * \ \frac{\langle k \rangle^k}{k!} \ = \ \boldsymbol{e^{-2.999}} \ * \ \boldsymbol{\frac{-2.999^{\,k}}{k!}}$ $(0 \ \leq \ k \ < \ N)$

3. Cây Cayley (Cayley tree)

Cây Cayley là cây đối xứng, được xây dựng bắt đầu từ nút trung tâm bậc k. Mỗi nút ở khoảng cách d tính từ nút trung tâm có bậc k, cho đến khi chúng ta đến các nút ở khoảng cách P có bậc một và được gọi là các lá. Ví dụ, hình 2 là cây Cayley có k = 3 và P = 5

a. Tính tổng số nút trên cây sau t bước tính từ từ nút trung tâm.

- Ta có thể thấy là nếu P = 1 thì sẽ tạo ra k node mới. Nhưng với P > 1 thì sẽ chỉ tạo ra k-1 node mới (1 liên kết sẽ nối trở lại parent node trước đó).
- Ta tính tổng số node tạo thành từ 1 đỉnh của P = 1 sau h bước (h = t - 1):

$$(k-1)^0 \ + \ (k-1)^1 \ + \ \dots \ + \ (k-1)^h \ = \ \frac{(k-1)^{h+1} - 1}{(k-1)-1} \ = \ \frac{(k-1)^t - 1}{(k-1)-1}$$

- Sau đó ta nhân số node được tạo thành từ P = 1, là bằng k (vì từ root node tạo thành) và cộng thêm 1 (root node):

$$\boldsymbol{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1}$$

- Với ví dụ cây Cayley có k=3 và P=5 thì tổng số node sau 5 bước là:

$$k * \frac{(k-1)^t - 1}{(k-1)-1} + 1 = 3 * \frac{(3-1)^5 - 1}{(3-1)-1} + 1 = 94$$

b. *Tính độ phân phối bậc (degree distribution) của mạng.*

- Cayley tree chỉ có 2 loại node: node lá (với bậc là 1), các node còn lại (với bậc là k)
- Số lượng node lá: $k(k-1)^{t-1}$

$$\Rightarrow P(1) = \frac{k(k-1)^{t-1}}{\text{Total nodes}} = \frac{\boldsymbol{k(k-1)^{t-1}}}{\boldsymbol{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1}}$$

- Số lượng các node không phải node lá: $k * \frac{(k-1)^t - 1}{(k-1)-1} + 1 - k(k-1)^{t-1}$

$$\Rightarrow P(k) = \frac{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1 - k(k-1)^{t-1}}{\text{Total nodes}} = \frac{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1 - k(k-1)^{t-1}}{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1}$$

$$= \boldsymbol{1 - \frac{k(k-1)^{t-1}}{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1}}$$

- Với ví dụ cây Cayley có k=3 và P=5 thì phân phối bạc của mạng là:
  - $P(1) = \frac{k(k-1)^{t-1}}{k * \frac{(k-1)^t - 1}{(k-1)-1} + 1} = \frac{3(3-1)^{5-1}}{94} = \frac{48}{94} \approx 0.5106$
  - $P(3) = 1 - P(1) \approx 0.4894$

c. *Tính đường kính d_{max}.*

- Đường kính $\mathbf{d_{max} = 2P}$
- Với ví dụ cây Cayley có k=3 và P=5 thì $d_{max} = 2P = 2*5 = 10$

d. *Tìm biểu thức của đường kính dmax theo tổng số nút N.*

- Từ câu a, ta có: $N = k * \frac{(k-1)^P - 1}{(k-1)-1} + 1$
- Do $d_{max} = 2P$, nên: $N = k * \frac{(k-1)^{\frac{d_{max}}{2}} - 1}{(k-1)-1} + 1$

$$\Leftrightarrow N - 1 = k * \frac{(k-1)^{\frac{d_{max}}{2}} - 1}{(k-1)-1}$$

$$\Leftrightarrow \frac{(N-1)(k-2)}{k} = (k-1)^{\frac{d_{max}}{2}} - 1$$

$$\Leftrightarrow \frac{(N-1)(k-2)}{\mathrm{k}} + 1 = (k-1)^{\frac{\mathrm{d_{max}}}{2}}$$

$$\Leftrightarrow \frac{\mathrm{d_{max}}}{2} = \log_{\mathrm{k-1}} \frac{(N-1)(k-2)}{\mathrm{k}}$$

$$\Leftrightarrow \boldsymbol{d_{max} = 2log_{k-1} \frac{(N-1)(k-2)}{k}}$$

4. Nghịch lý tình bạn (Friendship Paradox)

Phân phối bậc p_k là xác suất mà một nút được chọn ngẫu nhiên có k hàng xóm. Tuy nhiên, nếu chúng ta chọn ngẫu nhiên một liên kết, xác suất để một nút ở một trong các đầu của nó có bậc k là $q_k = Akp_k$, trong đó A là hệ số chuẩn hóa.

a. *Tìm hệ số chuẩn hóa A, giả sử rằng mạng có phân bố bậc theo luật mũ với $2 < \gamma < 3$, với bậc nhỏ nhất k_{min} và bậc lớn nhất k_{max}.*

- Để tìm hệ số chuẩn hóa A, ta dùng công thức:

$$\int_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{q_k dk} = 1$$

$$\Leftrightarrow \int_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{Akp_k dk} = 1$$

- Ta có: $\mathrm{p_k} = \mathrm{Ck}^{-\gamma}$(Sử dụng the normalization condition)

$$\Leftrightarrow \int_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{AkCk}^{-\gamma}\mathrm{dk} = 1$$

$$\Leftrightarrow \int_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{ACk}^{1-\gamma}\mathrm{dk} = 1$$

$$\Leftrightarrow \mathrm{A}\int_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{Ck}^{1-\gamma}\mathrm{dk} = 1$$

- Theo tài liệu, ta có công thức: $\mathrm{C} = (\gamma\text{-}1)\mathrm{k}_{\mathrm{min}}^{\gamma-1}$

$$\Leftrightarrow \mathbf{A} = \frac{\mathbf{1}}{\int_{k_{min}}^{k_{max}} (\boldsymbol{\gamma}-\mathbf{1})\boldsymbol{k_{min}^{\gamma-1}} * \boldsymbol{k^{1-\gamma}dk}}$$

b. *Chọn ngẫu nhiên một nút trong mạng có $N = 10^4$, γ= 2.3, k_{min}= 1 và k_{max}= 1000. Tính bậc trung bình của các nút lân cận.*

- Theo tài liệu, ta có công thức: $\langle \mathrm{k_n} \rangle = \sum_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{kq_k}$

$$\Leftrightarrow \langle \mathrm{k_n} \rangle = \sum_{\mathrm{k_{min}}}^{\mathrm{k_{max}}} \mathrm{ACk}^{2-\gamma}$$

$$\Leftrightarrow \langle \mathrm{k_n} \rangle = \mathrm{AC}^* \frac{\mathrm{k_{max}^{3-\gamma}}\text{-}\mathrm{k_{min}^{3-\gamma}}}{3\text{-}\gamma}$$