Network Science

2019. október 8.

Network models

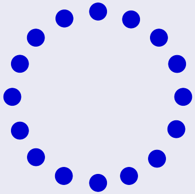

THE ERDŐS-RÉNYI MODEL

The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently of each other with probability p.



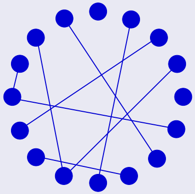
- This is also called as $G(N,p)$ model.
- The $G(N,M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently of each other with probability p.



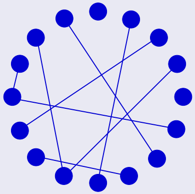
- This is also called as $G(N, p)$ model.
- The $G(N, M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently of each other with probability p.



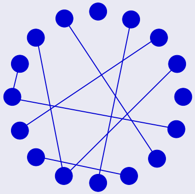
- This is also called as $G(N,p)$ model.
- The $G(N,M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

The Erdős-Rényi model (1959)

The Erdős-Rényi model (classical random graphs)

Pál Erdős and Alfréd Rényi:

- Take N nodes.
- Uniformly link each pair independently of each other with probability p.



- This is also called as $G(N,p)$ model.
- The $G(N,M)$ model is almost the same: distribute M links independently amongst the N nodes with uniform probability.

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property.

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property.
- average degree?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property.
- average degree: $\langle k \rangle = (N-1)p \simeq Np$,

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property.
- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links?

Properties of the Erdős-Rényi model

Properties of the E-R model:

- small world property.
- average degree: $\langle k \rangle = (N-1)p \simeq Np$,
- number of expected links: $M = pN(N-1)/2$.

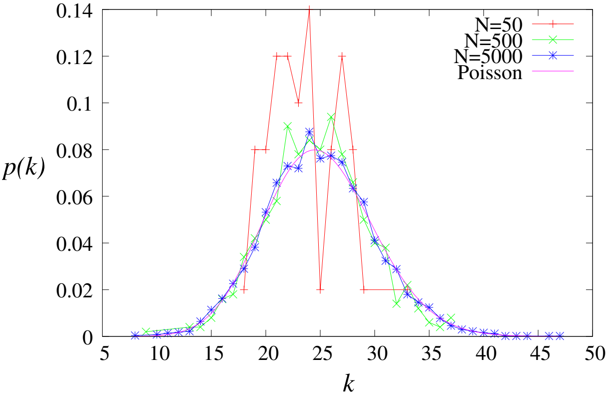
Degree distribution of the E-R graph

Degree distribution of the E-R graph

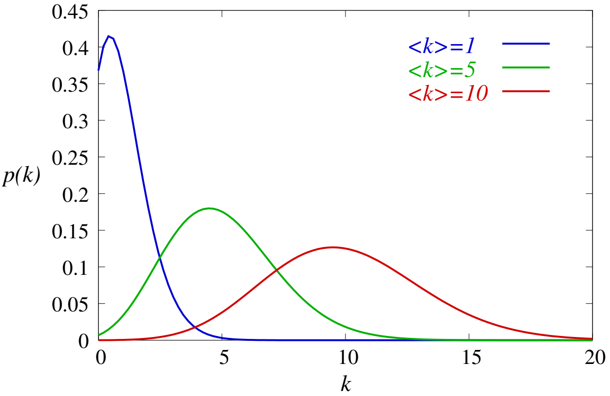
We have already derived the $p(k)$ of the E-R graph earlier,

$$\begin{aligned} p(k) &= \binom{N-1}{k} p^k (1-p)^{N-1-k} \simeq \binom{N}{k} p^k (1-p)^{N-k} && \text{(binomial)} \\ &\simeq \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} && \text{(Poisson)} \end{aligned}$$

Degree distribution of the E-R graph



Degree distribution of the E-R graph



Degree distribution of the E-R graph

Variance

- What is the variance of $p(k)$?

The degree distribution is binomial,

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k},$$

→ the average and variance for a binomial distribution in general is

$$\begin{aligned}
\langle k \rangle &= Np, \\
\langle k^2 \rangle &= Np(1-p) + p^2 N^2, \\
\mathrm{Var}(k) &= \langle k^2 \rangle - \langle k \rangle^2 = Np(1-p), \\
\sigma(k) &= \sqrt{\mathrm{Var}(k)} = \sqrt{Np(1-p)}.
\end{aligned}$$

Degree distribution of the E-R graph

Variance

- What is the variance of $p(k)$?

  The degree distribution is binomial,

$$p(k) = \binom{N}{k} p^k (1-p)^{N-k},$$

→ the average and variance for a binomial distribution in general is

$$\begin{aligned}
\langle k \rangle &= Np, \\
\langle k^2 \rangle &= Np(1-p) + p^2 N^2, \\
\mathrm{Var}(k) &= \langle k^2 \rangle - \langle k \rangle^2 = Np(1-p), \\
\sigma(k) &= \sqrt{\mathrm{Var}(k)} = \sqrt{Np(1-p)}.
\end{aligned}$$

Degree distribution of the E-R graph

Variance

- What happens in the $N \to \infty$ limit for a „realistic" E-R graph? (I.e., an E-R graph modeling a real system).

Degree distribution of the E-R graph
Variance

- What happens in the $N \to \infty$ limit for a „realistic” E-R graph? (I.e., an E-R graph modeling a real system).

It must remain sparse! $\to$ $\langle k \rangle$ =const.,

$$\left.\begin{array}{l} N \to \infty \\ \langle k \rangle = Np \to \text{const.} \end{array}\right\} \Rightarrow p \to 0$$

$$\mathrm{Var}(k) = Np(1-p) = \langle k \rangle (1-p)$$

$$\to \quad \mathrm{Var}(k) \to \langle k \rangle = \text{const.}$$

- What happens in the $N \to \infty$ limit for a „realistic” E-R graph? (I.e., an E-R graph modeling a real system).

It must remain sparse! $\to$ $\langle k \rangle$ =const.,

$$\left.\begin{array}{l} N \to \infty \\ \langle k \rangle = Np \to \text{const.} \end{array}\right\} \Rightarrow \ p \to 0$$

$$\mathrm{Var}(k) = Np(1-p) = \langle k \rangle (1-p)$$

$$\to \quad \mathrm{Var}(k) \to \langle k \rangle = \text{const.}$$

The variance is **constant**, thus, it becomes **negligible compared to the system size**!

- In the continuum formalism:

$$\mathcal{P}(k > k_0) = \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k\rangle^k}{k!} e^{-\langle k\rangle} dk.$$

- E.g, for $\langle k\rangle = 10$:
  - the prob. to find a node with $k \geq 20$ is 0.00158826,
  - the prob. to find a node with $k \leq 1$ is 0.00049,
  - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

→ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

→ A random society would be extremely homogeneous, with no outliers!

Degree distribution of the E-R graph

Variance

- In the continuum formalism:

$$\mathcal{P}(k > k_0) = \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k\rangle^k}{k!} e^{-\langle k\rangle} dk.$$

- E.g, for $\langle k\rangle = 10$:
  - the prob. to find a node with $k \geq 20$ is 0.00158826,
  - the prob. to find a node with $k \leq 1$ is 0.00049,
  - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

$\rightarrow$ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

$\rightarrow$ A random society would be extremely homogeneous, with no outliers!

Degree distribution of the E-R graph
Variance

- In the continuum formalism:

$$\mathcal{P}(k > k_0) = \int_{k_0}^{\infty} p(k)dk = \int_{k_0}^{\infty} \frac{\langle k \rangle^k}{k!} e^{-\langle k \rangle} dk.$$

- E.g, for $\langle k \rangle = 10$:
    - the prob. to find a node with $k \geq 20$ is 0.00158826,
    - the prob. to find a node with $k \leq 1$ is 0.00049,
    - the prob. to find a node with $k \geq 100$ is $1.79967152 \times 10^{-13}$.

- According to sociologists, for a typical individual $k \sim 1000$.

$\rightarrow$ the prob. to find someone with $k \geq 2000$ is roughly 10^{-27}!

$\rightarrow$ A random society would be extremely homogeneous, with no outliers!

Clustering coefficient in the E-R graph

- What is the clustering coefficient in the E-R graph?

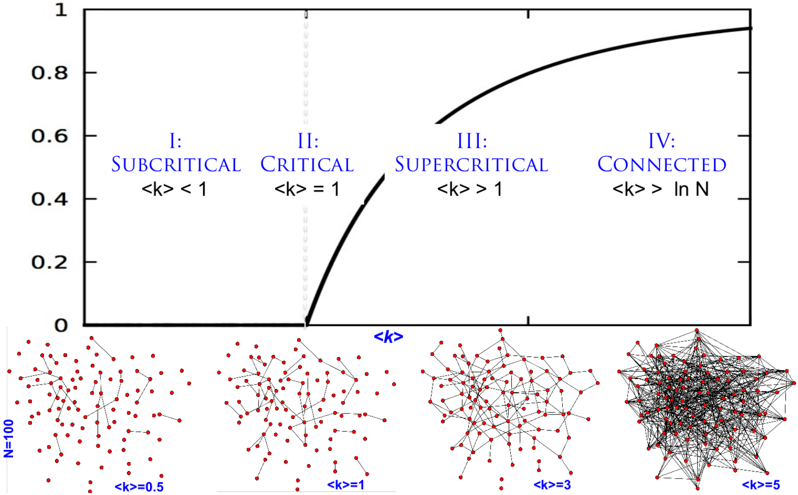
Clustering coefficient in the E-R graph

- What is the clustering coefficient in the E-R graph?
  - The E-R graph is very democratic, and we expect all nodes to have more or less the same C, thus, $C_i \simeq \langle C \rangle$.
  - C_i can be also interpreted as the probability of the neighbors of i being connected. Since in the E-R model we link every pair independently with uniform probability p, the neighbors of any node shall be linked also with probability p.
  - Thus, in the E-R graph $\langle C \rangle = p$.

Giant component in the E-R graph

Giant component in the E-R graph

Relative size of the largest component:



(from the slides of A.-L. Barabási)

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:

    - it is either non existent $\rightarrow$ probability$= 1 - p$,

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent $\rightarrow$ probability$= 1 - p$,
    - or j is also not in the giant component

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent $\rightarrow$ probability$= 1 - p$,
    - or j is also not in the giant component $\rightarrow$ probability$= pu$.

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent → probability$= 1 - p$,
    - or j is also not in the giant component → probability$= pu$.

→ The probability that i is NOT in the giant component is

$$(1 - p + pu)^{N-1}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.
  - Suppose node i is NOT in the giant component, and let us examine the connection between i and j:
    - it is either non existent $\rightarrow$ probability$= 1 - p$,
    - or j is also not in the giant component $\rightarrow$ probability$= pu$.

$\rightarrow$ The probability that i is NOT in the giant component is

$$(1 - p + pu)^{N-1}$$

- However, the probability that i is NOT in the giant component is u by definition, thus

$$u = (1 - p + pu)^{N-1}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component,** $S = \frac{s_1}{N}$**?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u = (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned} u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\ \ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned} u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\ \ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \\ \ln u &\approx -(N-1)\left[\frac{\langle k \rangle}{N-1}(1-u)\right] \end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$\begin{aligned}
u &= (1 - p(1-u))^{N-1} = \left(1 - \frac{\langle k \rangle}{N-1}(1-u)\right)^{N-1} \\
\ln u &= (N-1)\ln\left[1 - \frac{\langle k \rangle}{N-1}(1-u)\right] \\
\ln u &\approx -(N-1)\left[\frac{\langle k \rangle}{N-1}(1-u)\right] \\
u &\approx e^{-\langle k \rangle (1-u)}
\end{aligned}$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

Giant component in the E-R graph

- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

$$S \approx 1 - e^{-\langle k \rangle S}.$$

Giant component in the E-R graph

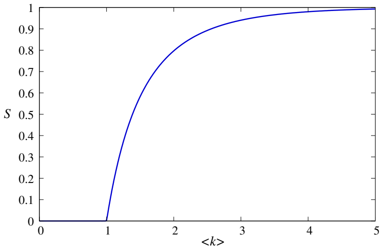
- **How to calculate the relative size of the largest component, $S = \frac{s_1}{N}$?**
  - Let $u = 1 - S$ denote the fraction of nodes NOT in the giant component.

$$u = (1 - p + pu)^{N-1}.$$

$$u \approx e^{-\langle k \rangle (1-u)}.$$

$$S \approx 1 - e^{-\langle k \rangle S}.$$

This equation can be solved numerically for any $\langle k \rangle$, giving the corresponding S:



Giant component in the E-R graph

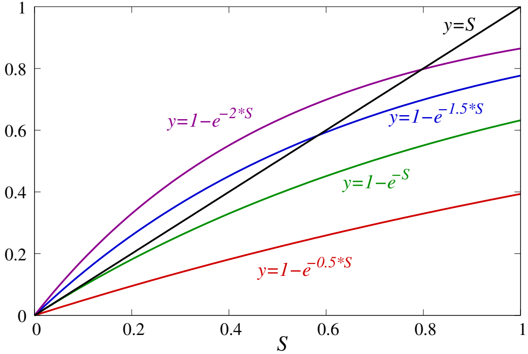
- **Where is the „critical” point?**

Giant component in the E-R graph

- **Where is the „critical" point?**
  - The critical point is where S becomes larger than zero.

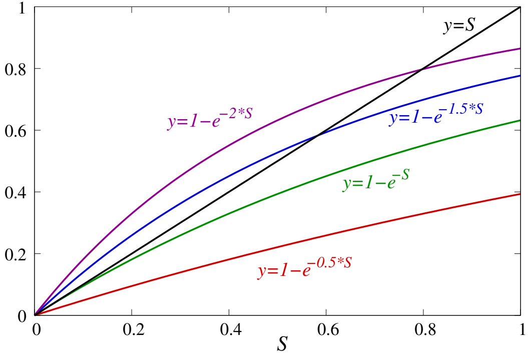
Giant component in the E-R graph

- **Where is the „critical" point?**
  - The critical point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ „graphically":



Giant component in the E-R graph

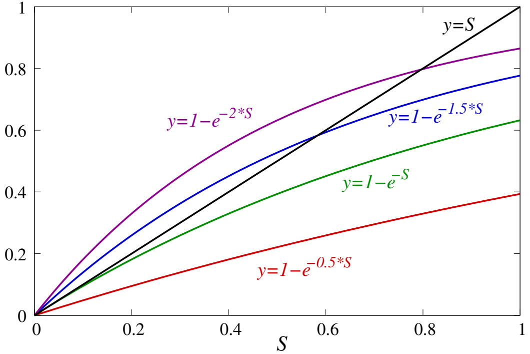
- **Where is the „critical” point?**
  - The critical point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ „graphically”:



→ What is the condition for having a non-trivial solution?

Giant component in the E-R graph

- **Where is the „critical” point?**
  - The critical point is where S becomes larger than zero.
  - Let us solve $S = 1 - e^{-\langle k \rangle S}$ „graphically”:



→ What is the condition for having a non-trivial solution?

$$\left.\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1\right|_{S=0}$$

Giant component in the E-R graph

- **Where is the „critical" point?**
  - The critical point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1 - e^{-\langle k \rangle S}\right) \geq 1 \Big|_{S=0}$$

Giant component in the E-R graph

- **Where is the „critical” point?**
  - The critical point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1-e^{-\langle k\rangle S}\right) \geq 1\Big|_{S=0}$$

$$\langle k\rangle\, e^{-\langle k\rangle S}\Big|_{S=0} \geq 1$$

Giant component in the E-R graph

- **Where is the „critical" point?**
  - The critical point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1-e^{-\langle k\rangle S}\right) \geq 1\Big|_{S=0}$$

$$\langle k\rangle e^{-\langle k\rangle S}\Big|_{S=0} \geq 1$$

$$\rightarrow \langle k\rangle \geq 1$$

Giant component in the E-R graph

- **Where is the „critical" point?**
  - The critical point is where S becomes larger than zero.

$$\frac{d}{dS}\left(1-e^{-\langle k\rangle S}\right)\geq 1\Big|_{S=0}$$

$$\langle k\rangle\, e^{-\langle k\rangle S}\Big|_{S=0}\geq 1$$

$$\rightarrow \langle k\rangle \geq 1$$

- **Thus, the critical point is at $k=1$, and for $k\geq 1$ we have a giant component in the E-R graph.**

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

(In fact the generating function formalism can be useful in many other problems as well...)

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

(In fact the generating function formalism can be useful in many other problems as well...)

- Our **main assumptions**:

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

(In fact the generating function formalism can be useful in many other problems as well...)

- Our **main assumptions**:
  - we know the degree distribution $p(k)$,

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

(In fact the generating function formalism can be useful in many other problems as well...)

- Our **main assumptions**:
  - we know the degree distribution $p(k)$,
  - we assume no degree correlations,

Percolation in a general random graph

- What if we would like to study the emergence (or disappearance) of the giant component in another type of random graph?

→ The **generating function formalism** provodes a general tool for studying that.

(In fact the generating function formalism can be useful in many other problems as well...)

- Our **main assumptions**:
  - we know the degree distribution $p(k)$,
  - we assume no degree correlations,
  - we approach the critical point of the percolation transition from below, i.e., from the dispersed phase where the network is sparse and can be assumed to be **localy tree-like**.

Percolation and generating functions

Quick overview on generating functions in general:

Percolation and generating functions

Quick overview on generating functions in general:

- For a discrete random variable X that can take non-negative integer values with a probability distribution given by $P(X=k)=p_X(k)$, the corresponding **generating function** is defined as

$$G_X(z)=\sum_{k=0}^{\infty} z^k p_X(k).$$

Percolation and generating functions

Quick overview on generating functions in general:

- For a discrete random variable X that can take non-negative integer values with a probability distribution given by $P(X = k) = p_X(k)$, the corresponding **generating function** is defined as

$$G_X(z) = \sum_{k=0}^{\infty} z^k p_X(k).$$

- The $p_X(k)$ probabilities defining the distributions can be obtained from the generating function as

$$p_X(k) = \frac{1}{k!} \left. \frac{dG_X(z)}{dz^k} \right|_{z=0} .$$