

Random graphs

Anne Bouillard

Introduction

Complex networks have received a lot of attention recently and led to the study of large graphs and the fundamental properties of real networks. Empirically, those real networks have four important properties:

- sparse: the degree of the vertices are very small compared with the size of the graph;
- scale-free: there are some vertices with high degree. For example, the distribution of the degrees is a power-law (for some $\tau > 1$, the number of vertices with degree k is proportional to $k^{-\tau}$);
- small world: the length between most of the vertices is relatively small;
- transitivity/clustering: the neighbors of my neighbors are my neighbors.

Examples of such graphs are social relations, the Internet, citation networks of scientists, telephony networks...

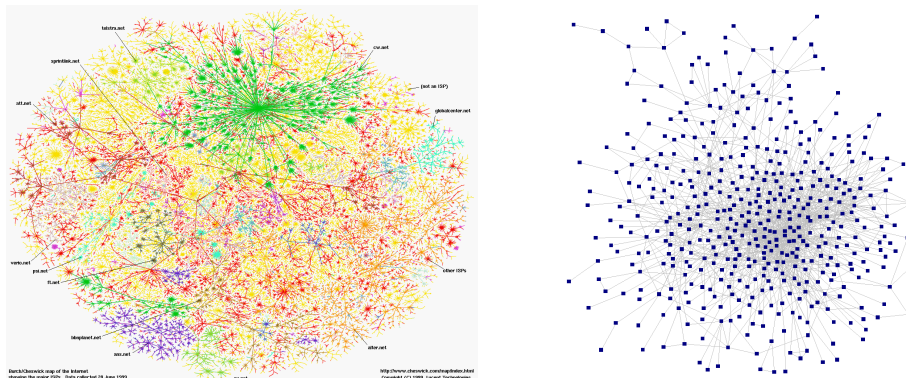


Figure 1: "Real world" topologies. Left: the Internet topology in 1999; right: collaboration graph of mathematicians in 2004.

To study such large graphs, *random graphs* plays an important role. Indeed, such graphs are described by local rules (for one vertex, who are its neighbors?) and possess, for some models, with high probability the properties describes above. As a consequence, it is believed that studying random graphs will help to understand the structure of large graphs. The simplest model of random graphs is the Eröds-Rényi graphs that have been developed in the late 1950s. In this model, every edge has the same probability p independently. Although this model is simplistic and does not exhibit the scale-free, power-law and small-world properties, they are worth being studied and are the starting point of the study of random graphs. Then, other models have been developed to take into account the properties of large real networks: preferential attachment, configuration models...

The aim of this course is to give an introduction to random graphs. We will mainly focus on Erdős-Rényi graphs, and exhibit interesting phenomena as the threshold functions and the emergence of the giant component. Then, we will study some complement, like preferential attachment and structures networks.

1 Notations and technical background

1.1 Probability theory

In this document, as we will only deal with non-negative integer random variables, the definitions and notations are defined for this particular context only.

Let X be a non-negative random variable.

- The expectation of X is $\mathbf{E}[X] = \sum_{i=0}^{\infty} i\mathbf{P}(X = i)$. This quantity is potentially $+\infty$.
- The variance of X is $\mathbf{Var}(X) = \mathbf{E}[(X - \mathbf{E}[X])^2]$.
- The (moment) generating function of X is $g_X(s) = \mathbf{E}[s^X]$. The generating function of a random variable characterizes its distribution.

The expectation has the following properties:

- *linearity*: $\forall a, b \in \mathbb{R}, \mathbf{E}[aX + bY] = a\mathbf{E}[X] + b\mathbf{E}[Y]$
- *monotony*: if $X \geq Y$, then $\mathbf{E}[X] \geq \mathbf{E}[Y]$
- if X and Y are independent, then $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$. As a consequence, $\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$ and $g_{X+Y} = g_X g_Y$.

1.1.1 Useful distributions

Definition 1. Let X be a non-negative integer random variable.

- X is distributed according to a binomial law with parameters n and p if

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

We write $X \sim \mathcal{Bin}(n, p)$.

- X is distributed according to a Poisson law with parameter λ if

$$\forall k \in \mathbb{N}, \quad \mathbf{P}(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

We write $X \sim \mathcal{Poi}(\lambda)$.

By abuse of the notation, in the following, $\mathcal{Bin}(n, p)$ will sometimes be directly used as a random variable with distribution $\mathcal{Bin}(n, p)$ independent from the rest of the model.

Theorem 1 (Limit of a binomial law). Let X_n be random variable such that $X_n \sim \mathcal{Bin}(n, p)$ and $np \xrightarrow{n \rightarrow \infty} \lambda$. Then

$$\lim_{n \rightarrow \infty} \mathbf{P}(X_n = k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

1.1.2 Useful inequalities

Lemma 1 (Markov inequality). *Let X be non-negative random variable with a finite expectation and $a > 0$. Then*

$$\mathbf{P}(X \geq a) \leq \frac{\mathbf{E}[X]}{a}.$$

Proof: Define the random variable $Y = a\mathbf{1}_{X \geq a}$. We have $Y \leq X$ (if $X < a$, $Y = 0 \leq X$ and if $X \geq a$ then $Y = a \leq X$). Then, by monotony of the expectation,

$$\mathbf{E}[X] \geq \mathbf{E}(Y) = a\mathbf{P}(X \geq a).$$

□

Corollary 1. *Let X be a non-negative integer variable. Then*

$$\mathbf{P}(X \neq 0) \leq \mathbf{E}[X].$$

Proof: $\mathbf{P}(X \neq 0) = \mathbf{P}(X \geq 1) \leq \mathbf{E}[X]/1 = \mathbf{E}[X]$.

□

Lemma 2 (Tchebychev inequality). *Let X be a random variable with finite expectation and variance. Then*

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq a) \leq \frac{\mathbf{Var}(X)}{a^2}.$$

Proof: $(X - \mathbf{E}[X])^2$ is a non-negative random variable with finite expectation $\mathbf{E}[(X - \mathbf{E}[X])^2] = \mathbf{Var}(X)$. Now,

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq a) = \mathbf{P}((X - \mathbf{E}[X])^2 \geq a^2) \leq \frac{\mathbf{E}[(X - \mathbf{E}[X])^2]}{a^2} = \frac{\mathbf{Var}(X)}{a^2}.$$

□

Corollary 2 (Second moment method). *Let X be a non-negative integer random variable. Then*

$$\mathbf{P}(X = 0) \leq \frac{\mathbf{Var}(X)}{\mathbf{E}[X]^2}.$$

Proof: $\mathbf{P}(X = 0) \leq \mathbf{P}(X \leq 0) \leq \mathbf{P}(X \leq 0 \text{ or } X \geq 2\mathbf{E}[X]) \leq \mathbf{P}(|X - \mathbf{E}[X]| \geq \mathbf{E}[X]) \leq \mathbf{Var}(X)/\mathbf{E}[X]^2$.

□

Lemma 3 (Chernoff bounds). *Let $X \sim \mathcal{Bin}(n, p)$, and $\mu = \mathbf{E}[X]$. Then for $\delta > 0$,*

$$\mathbf{P}(X \geq \mu(1 + \delta)) \leq e^{-\frac{\mu\delta^2}{3}}.$$

$$\mathbf{P}(X \leq \mu(1 - \delta)) \leq e^{-\frac{\mu\delta^2}{2}}.$$

2 Erdős-Rényi graphs

Let $n \in \mathbb{N}$ and $p \in [0, 1]$. The space $\mathcal{G}(n, p)$ is the space of undirected graphs with n vertices and where each edge has probability p independently from the others. More precisely, $\mathcal{G}(n, p) = (\Omega_n, \mathcal{P}(\Omega_n), P)$, where

- Ω_n is the set of non-directed graphs with n vertices $\{1, \dots, n\}$
- if for $1 \leq u < v \leq n$ $E_{u,v}$ is the event “there is an edge between u and v ”, $(E_{u,v})$ is a family of mutually independent events and $\mathbf{P}(E_{u,v}) = p$.

There are at most $N = \binom{n}{2}$ edges in a graph with n vertices and there are 2^N graphs in $\mathcal{G}(n, p)$. In the following, $G_{n,p}$ denotes a random element of $\mathcal{G}(n, p)$.

Example 1. In $\mathcal{G}(n, p)$,

- the complete graph has probability p^N ;
- the empty graph has probability $(1 - p)^N$;
- the probability that $G_{n,p}$ has m edges is $\binom{N}{m} p^m (1 - p)^{N-m}$.

Exercise 1

Erdős-Rényi model with n vertices and m edges

Originally, random graphs have been defined as $\mathcal{G}(n, m)$, which is the set of graphs with n vertices and m edges exactly. Graphs in $\mathcal{G}(n, m)$ are uniformly distributed.

1. Show that a graph in $\mathcal{G}(n, m)$ can be constructed as follows: starting from a graph with n vertices and no edge, choose one edge uniformly at random among the N possible edges. Add a second edge chosen uniformly at random from the $N - 1$ remaining edges and continue the same way until the graph has m edges.
2. Show that conditionally on having m edges, $G_{n,p}$ has the same distribution as in $\mathcal{G}(n, m)$.

Our goal here is to study the behavior of some graph properties when the number of vertices grows to infinity in two cases:

1. when p is fixed.
2. when $p = p(n)$ varies with n .

In the latter case, we are interested in finding *threshold functions*. A threshold function for the property A is a function $g(n)$ such that

- (i) if $\lim_{n \rightarrow \infty} p(n)/g(n) = 0$ (or $p \ll g$), then $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p(n)} \text{ has } A) = 0$.
- (ii) if $\lim_{n \rightarrow \infty} g(n)/p(n) = 0$ (or $p \gg g$), then $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p(n)} \text{ has } A) = 1$.

A threshold function can also be interpreted as follows: assign to each pair $\{u, v\}$ a random number $p_{u,v}$ chosen uniformly on $[0, 1]$. For $p \in [0, 1]$, the graph is made of the edges $\{u, v\}$ such that $p_{u,v} \leq p$. When p varies from 0 to 1, the graph $G_{n,p}$ grows. If $g(n) \gg p$, then $\mathbf{P}(G_{n,p} \text{ has } A) = 0$; and if $g(n) \ll p$, then $\mathbf{P}(G_{n,p} \text{ has } A) = 1$. Table 1 gives some examples of threshold functions.

Let us first focus on the first case, when p is fixed.

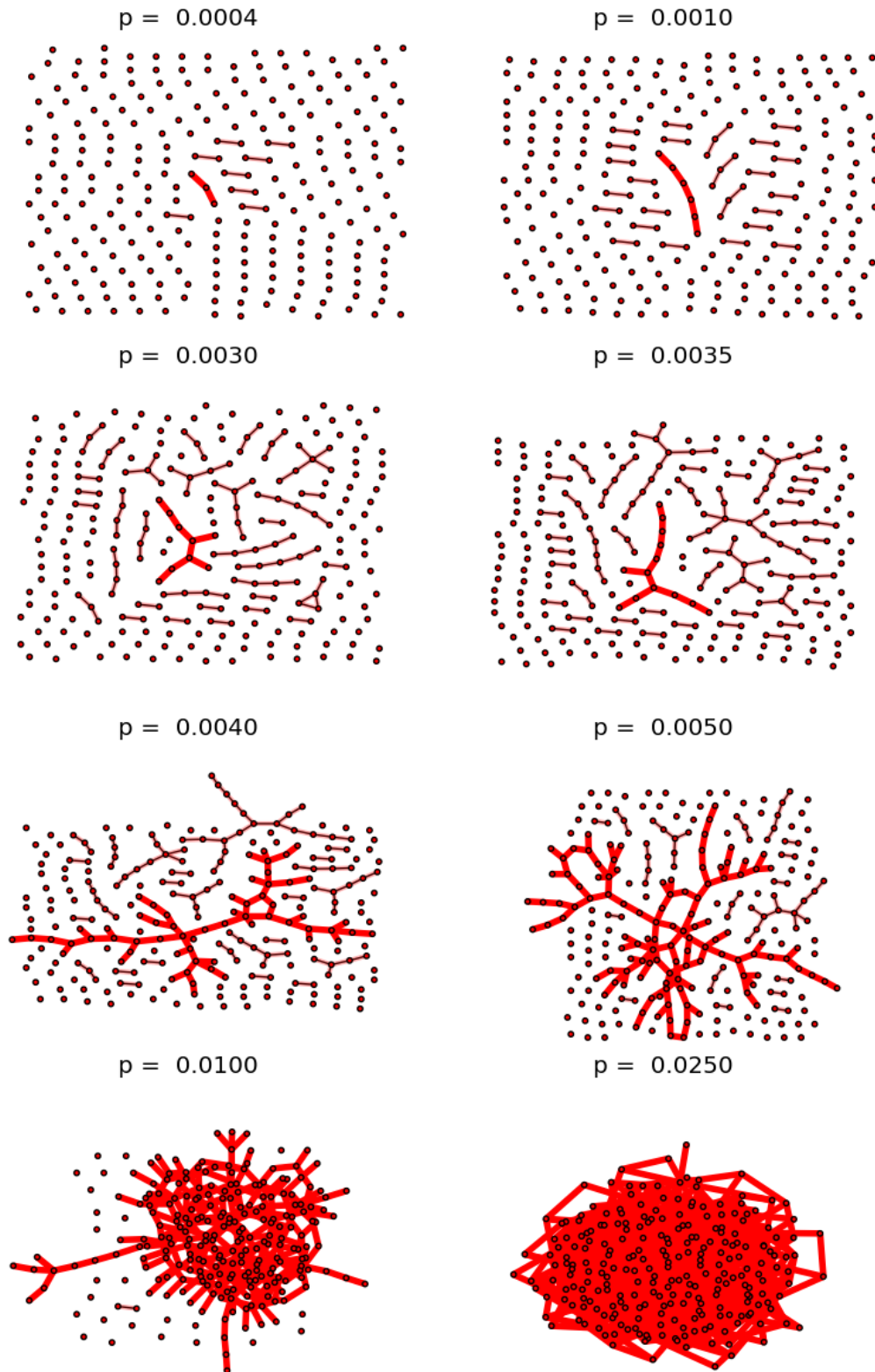


Figure 2: Example of Erdős-Rényi random graphs with 250 vertices and different probabilities.

property	threshold function $g(n)$
contains a path of length k	$n^{-\frac{k+1}{k}}$
is not planar	$\frac{1}{n}$
contains an Hamiltonian path	$\frac{\ln n}{n}$
is connected	$\frac{n}{\ln n}$
contains a clique of size k	$n^{-\frac{2}{k-1}}$

Table 1: Examples of threshold functions

2.1 Erdős-Rényi graphs and first order properties for a fixed p

We look at the closed form formulas generated by

$$F ::= \forall x F \mid \exists x F \mid F \vee F \mid F \wedge F \mid \neg F \mid x = y \mid I(x, y)$$

with the two axioms

$$\forall x \neg I(x, x) \quad \text{and} \quad \forall x \forall y I(x, y) \Leftrightarrow I(y, x).$$

Example 2. *The following properties are first-order:*

- *there exists a path of length 3:* $\exists x \exists y \exists z \exists w I(x, y) \wedge I(y, z) \wedge I(z, w)$;
- *there is no isolated vertex:* $\forall x \exists y I(x, y)$;
- *every triangle is included in a clique of size 4:* $\forall x \forall y \forall z (I(x, y) \wedge I(y, z) \wedge I(x, z) \Rightarrow \exists w (I(x, w) \wedge I(y, w) \wedge I(z, w))$.

The following properties are not first-order: G is connected, G is Hamiltonian, G is planar...

Theorem 2. *For every first-order statement A , $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ has } A) \in \{0, 1\}$.*

Proof: Let $A_{r,s}$ be the property $\forall x_1, \dots, x_r \forall y_1, \dots, y_s$ distinct vertices, $\exists z$ distinct vertex such that z is connected to every vertex x_i and none of y_j .

Fact 1. $\forall r, s, \lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ has } A_{r,s}) = 1$.

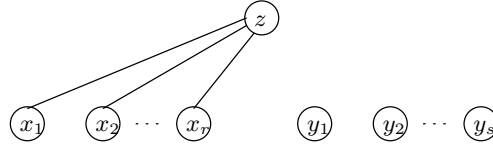
Let $A_{(x_i), (y_j), z}$ be the event “in $G_{n,p}$, z is connected to the vertices x_1, \dots, x_r and not to the vertices y_1, \dots, y_s ”. We have

$$\begin{aligned} \mathbf{P}(A_{(x_i), (y_j), z}) &= p^r (1-p)^s \\ \mathbf{P}(\forall z \neg A_{(x_i), (y_j), z}) &\leq (1 - p^r (1-p)^s)^n \\ \mathbf{P}(\exists (x_i), (y_j) \forall z \neg A_{(x_i), (y_j), z}) &\leq n^{r+s} (1 - p^r (1-p)^s)^n \\ \mathbf{P}(G_{n,p} \text{ has } A_{r,s}) &\leq 1 - n^{r+s} (1 - p^r (1-p)^s)^n \end{aligned}$$

Hence $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ has } A_{r,s}) = 1$.

To continue the proof, we need to use results from the completeness theory. We use the following results:

- If a system has a model, then it has a denumerable model.

Figure 3: z satisfies $A_{(x_i), (y_j), z}$.

- a theory T is complete if for all B , $T \cup B$ or $T \cup \neg B$ is inconsistent.

Let G and G' two graphs that satisfy $A_{r,s}$ for all s and r . Such graphs exist and can be constructed by induction:

1. G_0 is a graph with one vertex,
2. if G_n is built, then, for every disjoint subset of the vertices of G_n , S_1 and S_2 , either there exists a vertex in G_n that is adjacent to every vertex in S_1 and none in S_2 , or a new vertex satisfying that property is added to the graph. At the end of that step, the new graph obtained is G_{n+1} .

The limit of such graphs satisfies $A_{r,s}$ for all s and r . The graphs G_n are finite for all n but obviously the graph obtained as a limit is not finite. It is then countable and we can assume that G and G' have an infinite countable number of vertices.

Fact 2. G and G' are isomorphic.

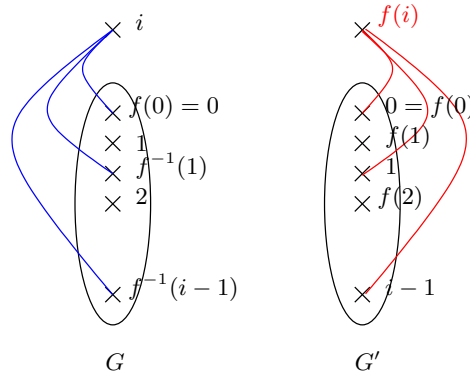
The set of vertices of G and G' is \mathbb{N} . We build an isomorphism by induction. Let f be this isomorphism and initially set $f(0) = 0$. Suppose that $f(0), f^{-1}(0), \dots, f(i-1), f^{-1}(i-1)$ have been defined. We now define $f(i)$ and $f^{-1}(i)$. Let

$$V = \{0, \dots, i-1, f^{-1}(0), \dots, f^{-1}(i-1)\}$$

be the set of vertices where f is already defined. Set $R = \{j \in V \mid \{j, i\} \text{ is an edge in } G\}$ and $S = \{j \in V \mid \{j, i\} \text{ is not an edge in } G\}$. From our hypothesis, there exists a vertex k in G' such that k is adjacent (in G') to every vertex in $f(R)$ and none in $f(S)$. Set $f(i) = k$ and $f^{-1}(k) = i$. As a consequence,

$$(i, j) \text{ is a edge in } G \Leftrightarrow (f(i), f(j)) \text{ is an edge in } G',$$

and the two graphs are isomorphic.

Figure 4: Construction of an isomorphism between G and G' .

Fact 3. *The system composed of all the $A_{r,s}$ is complete: for every first order statement B , either B or $\neg B$ is provable from the $(A_{r,s})$.*

By contradiction: suppose that both B and $\neg B$ are not provable. Then, the theories $(A_{r,s}) + B$ and $(A_{r,s}) + \neg B$ are both consistent and there exist models G and G' for both of them. But, from the previous fact, G and G' are isomorphic. Consequently, they cannot disagree on B .

To conclude, let A be a first order statement and suppose that A is provable from the $(A_{r,s})$. As proofs are finite, then A is provable from a finite set \mathcal{S} of $A_{r,s}$. Then,

$$\mathbf{P}(\neg A \text{ in } G_{n,p}) \leq \sum_{(r,s) \in \mathcal{S}} \mathbf{P}(\neg A_{r,s} \text{ in } G_{n,p}) \xrightarrow{n \rightarrow \infty} 0.$$

Then $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ has } A) = 1$. If A is not provable from the $A_{r,s}$, then the same holds for $\neg A$ and $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ has } A) = 0$, which ends the proof. \square

2.2 Phase transition in Erdős-Rényi graphs

2.2.1 A first (and simple) example

Theorem 3. *If A is "having a clique of size 4", then the threshold function is $g(n) = n^{-2/3}$. More precisely,*

- *if $p(n) \ll n^{-2/3}$, then $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ satisfies } A) = 0$;*
- *if $p(n) \gg n^{-2/3}$, then $\lim_{n \rightarrow \infty} \mathbf{P}(G_{n,p} \text{ satisfies } A) = 1$.*

Proof: The first assertion is proved using Markov inequality, and the second using the second moment method.

Let $C_1, \dots, C_{\binom{n}{4}}$ be an enumeration of the 4-vertex sets and define the random variables $X_i \in \{0, 1\}$, $i \in \{1, \dots, \binom{n}{4}\}$

$$X_i = 1 \Leftrightarrow C_i \text{ is a clique of size 4.}$$

Let $X = \sum_i X_i$.

- $\mathbf{E}[X] = \sum \mathbf{E}[X_i] = \binom{n}{4} p(n)^6 = (\frac{1}{24} n^4 + o(n^4)) p(n)^6$;
- $\mathbf{E}[X^2] = \sum \mathbf{E}[X_i] + \sum_{i \neq j} \mathbf{E}[X_i X_j]$. We need to consider several cases, depending of the number of common vertices in C_i and C_j . The case disjunction is shown in 2.

$ C_i \cap C_j $	$\mathbf{E}[X_i X_j]$	number
≤ 1	$p(n)^{12}$	$\binom{n}{4} ((\binom{n-4}{4}) + 4 \binom{n-4}{3})$
2	$p(n)^{11}$	$\binom{n}{4} 6 \binom{n-4}{2}$
3	$p(n)^9$	$\binom{n}{4} 4(n-4)$

Table 2: 4-cliques: case disjunction for $\mathbf{Var}(X)$.

Hence,

$$\mathbf{E}[X^2] = (\frac{1}{24} n^4 + o(n^4)) p(n)^6 + (\frac{1}{242} n^8 + o(n^8)) p(n)^{12} + (\frac{6}{242} n^6 + o(n^6)) p(n)^{11} + (\frac{4}{24} n^5 + o(n^5)) p(n)^9$$

and

$$\text{Var}[X] = (\frac{1}{24}n^4 + o(n^4))p(n)^6 + (o(n^8))p(n)^{12} + (\frac{6}{24^2}n^6)p(n)^{11} + (\frac{4}{24}n^5)p(n)^9.$$

Now,

- if $p(n) = o(n^{-2/3})$, then by the Markov inequality,

$$\mathbf{P}(X \neq 0) \leq \mathbf{E}[X] = (\frac{1}{24}n^4 + o(n^4))p(n)^6 = o(1).$$

- if $n^{-2/3} = o(p(n))$, $n^4p(n)^6 \xrightarrow{n \rightarrow \infty} \infty$ then by the second moment method,

$$\mathbf{P}(X = 0) \leq \frac{\text{Var}(X)}{\mathbf{E}[X]^2} = O(n^{-4}p(n)^{-6}) + o(1) + O(n^{-2}p(n)^{-1}) + O(n^{-3}p(n)^{-3}) = o(1).$$

□

Exercise 2

Conditional expectation inequality

1. (*Jensen inequality*) Let X be a real random variable, I an interval and $\phi : I \rightarrow \mathbb{R}$ a convex function, such that $\mathbf{P}(X \in I) = 1$. Show that if X and $\phi(X)$ are integrable, then $\mathbf{E}[\phi(X)] \geq \phi(\mathbf{E}[X])$.

Soit $X = \sum_{i=1}^n X_i$, où les X_i sont des variables aléatoires à valeurs dans $\{0, 1\}$. On veut montrer que

$$P(X > 0) \geq \sum_{i=1}^n \frac{P(X_i = 1)}{E(X | X_i = 1)}.$$

Soit $Y = 1/X$ si $X \neq 0$ et $Y = 0$ sinon.

2. Show that $P(X > 0) = E(XY)$.

3. Show that $E(X_i Y) \geq \frac{P(X_i=1)}{E(X | X_i=1)}$.

4. Conclude.

Exercise 3

Number of triangles in a graph

Consider a graph of $G_{n,p}$ with $p = 1/n$. Let X be its number of triangles.

1. Show that $P(X \geq 1) \leq 1/6$.

2. Show that $\lim_{n \rightarrow \infty} P(X \geq 1) \geq 1/7$. *Indication* : Use the previous exercise

2.2.2 Isolated vertices and connectivity

For a vertex x , define the random variable

$$I(x) = \begin{cases} 1 & \text{if } x \text{ is isolated} \\ 0 & \text{otherwise.} \end{cases}$$

Set

1. $I = \sum_x I(x)$ the number of isolated vertices,
2. $C = 1$ if and only if $G_{n,p}$ is connected.

We first deal with isolated vertices, but the threshold function is the same: when there is no isolated vertex, then with high probability, the graph will be connected.

Theorem 4. *If $pn - \ln n \rightarrow \infty$, then $G_{n,p}$ is connected with high probability and if $pn - \ln n \rightarrow \infty$, then the $G_{n,p}$ is disconnected with high probability.*

Exercise 4

The aim of this exercise is to prove Theorem 4.

1. Show that if $pn - \ln n \rightarrow +\infty$ then $\lim_{n \rightarrow \infty} \mathbf{P}(I \neq 0) = 0$.

We now assume that $pn - \ln n \rightarrow -\infty$.

2. Compute $\text{Var}(I)$ and show that $\text{Var}(I) \leq \mathbf{E}[I] + \mathbf{E}[I^2] \frac{p}{1-p}$.
3. Show that $\lim_{n \rightarrow \infty} \mathbf{P}(I = 0) = 0$.
4. Show that in this case $\lim_{n \rightarrow \infty} \mathbf{P}(C = 1) = 0$.

Now, let us deal with the connectivity above the threshold and compute the probability that there is no isolated vertex, but the graph is disconnected: $\mathbf{P}(C = 0, I = 0)$. Let X_k be the number of spanning tree of size k in the components of size k .

5. Show that $\mathbf{P}(C = 0, I = 0) \leq \sum_{k=2}^{n/2} \mathbf{E}[X_k]$.
6. Give an upper bound for $\mathbf{E}[X_k]$.

We now investigate the case where $p = a \ln n / n$ with $a > 1/2$. This case will be sufficient to study the case where $pn - \ln n \rightarrow +\infty$.

7. Show that when k is fixed, $\mathbf{E}[X_k] = o(1)$.
8. Using that $k(n-k) \geq kn/2$ and $x \mapsto xe^{-x/2}$ is decreasing for $x > 2$, show that $\mathbf{E}[X_k] \leq n^{1-k/4}$ for n large enough.
9. Conclude by showing that $\mathbf{P}(C = 0, I = 0) \xrightarrow{n \rightarrow \infty} 0$ when $pn - \ln n \rightarrow +\infty$.

3 Moment generating functions

Definition 2. Let X be a random variable on \mathcal{N} . Its (moment) generating functions is

$$g_X : s \mapsto \mathbf{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbf{P}(X = k).$$

g_X is \mathcal{C}^∞ on $] -1, 1[$. We have $g_X(0) = \mathbf{P}(X = 0)$, $g_X(1) = 1$, $\mathbf{P}(X = n) = g_X^{(n)}(0)/n!$, $\mathbf{E}[X] = g_X'(1)$.

Proposition 1. Let X and Y be two independent random variables, with respective generating functions g_X and g_Y . Then the generating function of $X + Y$ is $g_{X+Y} = g_X g_Y$.

Proof: For all s , $g_{X+Y}(s) = \mathbf{E}[s^{X+Y}] = \mathbf{E}[s^X] \mathbf{E}[s^Y]$. □

Example 3. • $X \sim \mathcal{Ber}(p)$: $g_X(s) = 1 - p + ps$;

• $X \sim \mathcal{Bin}(n, p)$: $g_X(s) = (1 - p + ps)^n$;

• $X \sim \mathcal{Poi}(\lambda)$: $g_X(s) = e^{\lambda(s-1)}$;

Proposition 2. Let X and Y be two random variables, with respective generating functions g_X and g_Y . If $\forall s \in [0, \delta]$, $g_X(s) = g_Y(s)$, then X and Y have the same distribution.

Theorem 5. Let T be a non-negative integer random variable and $(Z_i)_{i \in \mathbb{N}}$ be a sequence if i.i.d r.v. independent of T . Set $X = \sum_{i=0}^T Z_i$ and let g_Z , g_T and g_X be the respective generating functions of Z_1 , T and X . Then

$$g_X = g_T \circ g_Z.$$

Proof:

$$s^{Z_1 + \dots + Z_T} = \sum_{n=0}^{\infty} \mathbf{1}_{\{T=n\}} s^{Z_1 + \dots + Z_n},$$

so

$$\begin{aligned} \mathbf{E}(s^{Z_1 + \dots + Z_T}) &= \sum_{n=0}^{\infty} \mathbf{E}[\mathbf{1}_{\{T=n\}} s^{Z_1 + \dots + Z_n}] \text{ (linearity)} \\ &= \sum_{n=0}^{\infty} \mathbf{E}[\mathbf{1}_{\{T=n\}}] \mathbf{E}[s^{Z_1 + \dots + Z_n}] \text{ (independence of } T \text{ and } Z_i) \\ &= \sum_{n=0}^{\infty} \mathbf{P}(T = n) [g_Z(s)]^n \text{ (independence of the } Z_i) \\ &= \mathbf{E}[g_Z(s)^T] = g_T(g_Z(s)). \end{aligned}$$

□

Corollary 3 (Wald's equality). Let T be a non-negative integer random variable and $(Z_i)_{i \in \mathbb{N}}$ be a sequence if i.i.d r.v. independent of T . Let $X = \sum_{i=0}^T Z_i$. Let g_Z , g_T and g_X be the respective generating functions of Z_1 , T and X . Then

$$\mathbf{E}[X] = \mathbf{E}[Z] \mathbf{E}[T].$$

Proof:

$$\mathbf{E}[X] = g_X'(1) = g_Z'(1) g_T'(g_Z(1)) = g_Z'(1) g_T'(1) = \mathbf{E}[Z] \mathbf{E}[T].$$

□

3.1 Chernoff bounds

The idea of the Chernoff bounds is to apply Markov inequality to the generating function.

Theorem 6. • $\forall s > 1, P(X \geq a) \leq \inf_{s>1} \frac{\mathbf{E}(s^X)}{s^a}$

• $\forall s < 1, P(X \leq a) \leq \inf_{s<1} \frac{\mathbf{E}(s^X)}{s^a}$

Proof: $\forall s > 1, P(X \geq a) = P(s^X \geq s^a) \leq \frac{\mathbf{E}(s^X)}{s^a}$

$\forall s < 1, P(X \leq a) = P(s^X \geq s^a) \leq \frac{\mathbf{E}(s^X)}{s^a}$ □

Some useful particular cases:

Theorem 7. Let X_1, \dots, X_n be n independent r.v., $X_i \sim \text{Ber}(p_i)$. Let $X = \sum_{i=1}^n X_i$ and set $\mu = \mathbf{E}[X]$. Then

1. $\forall \delta > 0, P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right)^\mu$.

2. $\forall \delta \in]0, 1], P(X \geq (1 + \delta)\mu) \leq e^{-\mu \frac{\delta^2}{3}}$.

Proof: Let g_i the generating function of X_i . we have

$$g_i(s) = 1 - p_i + p_i s = 1 + p_i(s - 1) \leq e^{p_i(s-1)}.$$

As a consequence

$$g_X(s) = \prod_{i=1}^n g_i(s) \leq \prod_{i=1}^n e^{p_i(s-1)} = e^{\mu(s-1)}.$$

But, $\forall s > 1, P(X \geq (1 + \delta)\mu) \leq \frac{\mathbf{E}(s^X)}{s^{(1+\delta)\mu}} \leq \frac{e^{\mu(s-1)}}{s^{(1+\delta)\mu}}$. Now take $s = 1 + \delta$, we take

$$P(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right)^\mu.$$

To prove the other inequality, we just need to notice that

$$\forall \delta \in]0, 1], \frac{e^\delta}{(1 + \delta)^{1 + \delta}} = e^{\delta - (1 + \delta) \ln(1 + \delta)} \leq e^{-\frac{\delta^2}{3}}.$$

□

The following theorem is very similar:

Theorem 8. Let X_1, \dots, X_n be n independent r.v., $X_i \sim \text{Ber}(p_i)$. Let $X = \sum_{i=1}^n X_i$ and set $\mu = \mathbf{E}[X]$. Then for all $\delta \in]0, 1[$,

1. $P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1 - \delta}} \right)^\mu$.

2. $P(X \leq (1 - \delta)\mu) \leq e^{-\mu \frac{\delta^2}{2}}$.

Proof: The proof is exactly the same with $s < 1$:

$$P(X \leq (1 - \delta)\mu) \leq \frac{\mathbf{E}(s^X)}{s^{(1-\delta)\mu}} \leq \frac{e^{\mu(s-1)}}{s^{(1-\delta)\mu}}.$$

We choose $s = 1 - \delta$ and

$$P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^\mu.$$

To prove the other inequality, we just need to notice that

$$\forall \delta \in]0, 1[, \frac{e^\delta}{(1 - \delta)^{1-\delta}} = e^{\delta - (1-\delta) \ln(1-\delta)} \leq e^{-\frac{\delta^2}{2}}.$$

□

3.2 Galton-Watson branching process

The Galton-Watson branching process was initially introduced to study the extinction of family names in the Victorian England. The construction is the following:

- $X_0 = 1$ (the root, level 0)
- X_n is the number of nodes at level n (or at the n -th generation)

We denote by $Z_i^{(n)}$ the number of children of the i -th node of the n -th generation, and $(Z_i^{(n)})_{i,n}$ are i.i.d. with the same law as Z .

We have

$$X_{n+1} = \sum_{i=1}^{X_n} Z_i^{(n)}.$$

The simplest way to study this process is to use the moment generating functions. Set $g(s) = \mathbf{E}[s^Z]$ the generating function of Z , and $\phi_n = \mathbf{E}[s^{X_n}]$ that of X_n .

Lemma 4. $\phi_{n+1} = g_Z(\phi_n)$.

Proof: From the Wald equality, we have $\phi_{n+1} = \phi_n \circ g_Z$. Then,

$$\phi_{n+1} = \phi_0 \circ g_Z \circ \cdots \circ g_Z = \phi_0 \circ g_Z^{n+1}.$$

But $\mathbf{P}(X_0 = 1) = 1$, so $\phi_0(s) = s$ and $\phi_{n+1} = g_Z^{n+1}$. □

Let $p_e = \mathbf{P}(\exists n \in \mathbb{N}, X_n = 0) = \mathbf{P}(\cup_{n \in \mathbb{N}} \{X_n = 0\})$ the extinction probability of the process. As $\{X_n = 0\} \subseteq \{X_{n+1} = 0\}$, we have $p_e = \lim_{n \rightarrow \infty} \mathbf{P}(X_n = 0)$.

Lemma 5. $p_e = g_Z(p_e)$.

Proof: We know that $\phi_{n+1}(0) = g_Z(\phi_n(0))$. But $\phi_{n+1}(0) = \mathbf{P}(X_{n+1} = 0)$ and $\phi_n(0) = \mathbf{P}(X_n = 0)$. Then, by continuity (g_Z is continuous in 0), $p_e = g_Z(p_e)$. □

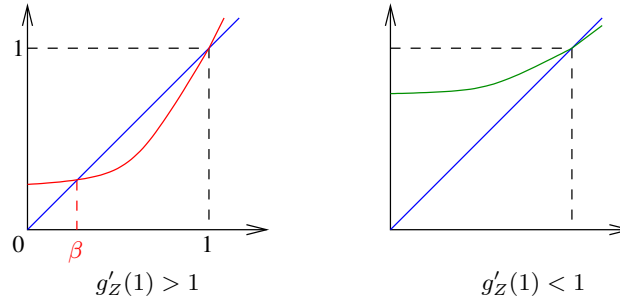
Theorem 9 (fixed point). *Consider the equation $p = g(p)$ where g is the generating function of a random variable X .*

1. g is non-decreasing and convex on $[0, 1]$. Moreover, if $\mathbf{P}(X = 0) < 1$, then g is strictly increasing, and if $\mathbf{P}(X \leq 1) < 1$, then g is strictly convex.
2. If $\mathbf{P}(X < 1) < 1$, and if $\mathbf{E}[X] \leq 1$, then the equation $x = g(x)$ has a unique solution in $[0, 1]$, $x = 1$. If $\mathbf{E}[X] > 1$, then the equation $x = g(x)$ has two solutions, in $[0, 1]$, $x = 1$ and $\beta \in [0, 1[$.

Proof:

1. $g_Z(s) = \sum_{n \in \mathbb{N}} \mathbf{P}(Z = n)s^n$ is non-decreasing and strictly increasing if $\mathbf{P}(Z = 0) < 1$. $g'_Z(s) = \sum_{n \in \mathbb{N}} \mathbf{P}(Z = n+1)s^n$ is non-decreasing and strictly increasing if $\mathbf{P}(Z \leq 1) < 1$, so g_Z is convex and strictly convex if $\mathbf{P}(Z \leq 1) < 1$.
2. $x = 1$ is trivially a solution. Now, we use the convexity of g_Z . If $\mathbf{E}[X] \leq 1$, then $g'_Z(1) \leq 1$ and, as the function is convex, $\forall x < 1, g'_Z(x) \leq 1$ and $g_Z(x) > x$.
If $\mathbf{E}[X] > 1$, on an interval $[1 - \epsilon, 1[$, $g_Z(x) < x$. But $g_Z(0) \geq 0$, so there exists β such that $\beta = g_Z(\beta)$.

□



Theorem 10. Let p_e be the extinction probability of the Galton-Watson process.

1. If $\mathbf{P}(Z > 1) > 0$ and $\mathbf{E}[Z] \leq 1$ then $p_e = 1$;
2. If $\mathbf{P}(Z > 1) = 0$ and $\mathbf{E}[Z] = 1$, then $p_e = 0$;
3. If $\mathbf{E}[Z] > 1$, then $p_e = \beta < 1$.

Proof: Let $x_n = \mathbf{P}(X_n = 0)$. We know that $x_0 = 0$, so $\beta - x_0 \geq 0$. Now, if $x_n \leq \beta$, then as g_Z is non-decreasing, $x_{n+1} = g_Z(x_n) \leq g_Z(\beta) = \beta$. So $p_e \leq \beta$ and finally $p_e = \beta$. □

Exercise 5

Total population

Let \mathcal{Z} be the total population of the branching process whose probability of extinction is 1, given by $\mathcal{Z} = \sum_{i,n} Z_i^{(n)}$. Let g_Z be its generating function.

Show that $g_Z(s) = sg_Z(g_Z(s))$.

Poisson branching process

Theorem 11. *If $Z \sim \text{Poi}(c)$, then*

1. *either $c \leq 1$, and $\mathbf{P}(T < \infty) = 1$.*
2. *or $c > 1$, and $\mathbf{P}(T = \infty)$ is the unique positive solution of the equation $z = e^{c(z-1)}$.*

Proof: For a variable that has a Poisson distribution, $\mathbf{E}[Z] = c$. So if $c \geq 1$, the extinction probability is 1. If $c > 1$, $g_Z(s) = e^{c(s-1)}$. \square

An alternative presentation of Galton-Watson branching processes Let Z be a r.v. on the non-negative integers with mean $\mathbf{E}[z] = c$.

- At time $t = 0$, the tree is only made of the root, which is numbered 1.
- At time 1, Z_1 children of the root are added to the tree, where $Z_1 \sim Z$, and are numbered $2, \dots, Z_1 + 1$.
- More generally, at time t , the node t is selected and Z_t children are added to node t , numbered from $2 + \sum_{i=1}^{t-1} Z_i$ to $1 + \sum_{i=1}^{t-1} Z_i + Z_t$, where $Z_t \sim Z$ and Z_t is independent from Z_1, \dots, Z_{t-1} . If at time t there is no node numbered t , then the process stops. At time t , the nodes $1, \dots, t-1$ are the dead nodes and the other are the living nodes.

Let Y_t be the number of living nodes at time t , then $Y_0 = 1$ and for $t > 0$, $Y_t = Y_{t-1} + Z_t - 1$. The process stops when $Y_t = 0$, but the variable Y_t can be defined even after the process stops.

- If for all $t \leq 0$, $Y_t > 0$, then the process does not stop and we set $T = \infty$
- if there exists $t \geq 0$ such that $Y_t = 0$, then T is the least integer such that $Y_t = 0$. The process stops at time T and T is the size of the process.

Exercise 6

Branching process conditionned on extinction

The *history* of a process is given by the sequence $H = \{Z_1, Z_1, \dots, Z_T\}$ of the number of children in a one-by-one exploration: for all $t < T$, $Y_t > 0$ and $Y_T = 0$.

1. Consider x_1, \dots, x_T a finite history. Express $\mathbf{P}(H = (x_1, \dots, x_k))$ in function of the distribution of Z .

In the remaining, Z has a Poisson distribution with parameter $\lambda > 1$. As a consequence, its extinction probability is $p_{ext} < 1$. Let $\phi(s) = e^{s-1}$.

Define $\mu = \lambda p_{ext}$.

2. Show that μ is the only solution of $\frac{\phi(\lambda)}{\lambda} = \frac{\phi(s)}{s}$ and $s < 1$.
3. Show that conditioned on extinction, the distribution of the histories coincides with the distribution of the histories under a Poisson offspring distribution with parameter μ .

Exercise 7*Branching processes in continuous time*

Consider the following process:

- at time 0, $Z_0 = 1$ (the root of the process). By convention, this node is born at time 0.
- when a node i is born, its lifetime has a exponential distribution with parameter μ : if it is born at time t , it dies at time $t + U_i$, with U_i exponentially distributed with parameter μ .
- a live node i can give birth to children. Children are generated according to an exponentially distribution with parameter λ : if a node is born at time t , its first child (if it is not dead before) is generated at time $t + V_i^{(1)}$, the second child at time $t + V_i^{(1)} + V_i^{(2)}$, and so on where $V_i^{(j)}$ is exponentially distributed with parameter λ .
- all the lifetimes (U_i) and birth interval $(V_i^{(j)})$ for a mutually independent family of random variables.

We recall that for X an exponentially distributed random variable with parameter λ satisfies: $\forall t \geq 0, \mathbf{P}(X \geq t) \leq e^{-\lambda t}$.

1. Show that the exponential distribution is memory-less: if X is exponentially distributed with parameter λ , $\forall t, u \geq 0$,

$$\mathbf{P}(X \geq t + u \mid X \geq t) = \mathbf{P}(X \geq u).$$

Let X_1 and X_2 be two independent exponentially distributed random variables with respective parameters λ and μ .

2. Show that $\min(X_1, X_2)$ is also exponentially distributed. What is its parameter?
3. What is the probability that $\min(X_1, X_2) = X_1$?

We are back to the branching process.

4. What is the law of the number of children for each node?
5. What is the probability of extinction of this process?

3.3 Emergence of cycles

Galton-Watson branching processes are used to study random graphs, specially when the average degree is small: in that case, with high probability, the graph structure is a forest, and then, locally, the connected components of the graph can be compared with branching processes. To illustrate this fact, let us focus on the emergence threshold of cycles in a Erdős-Rényi graph.

Let us denote by C the number of cycles in an Erdős-Rényi graph with n vertices. For $k \geq 3$. The number of potential cycles of length k in a random graph with n vertices is $\frac{n(n-1)\cdots(n-k+1)}{2k}$ (take k ordered vertices, divide by k for the starting point and 2 for the orientation). Then,

$$\begin{aligned} \mathbf{E}[C] &= \sum_{k=3}^n \frac{n(n-1)\cdots(n-k+1)p^k}{2k} \\ &\leq \sum_{k \geq 3} \frac{(np)^k}{2k} \leq \sum_{k \geq 3} (np)^k \\ &\leq \frac{(np)^3}{1 - np}. \end{aligned}$$

As a consequence,

- if $p = o(1/n)$, $\mathbf{P}(C > 0) \leq \frac{\mathbf{E}[C]}{1} \xrightarrow{n \rightarrow \infty} 0$ and
- if $p = c/n$ with $c < 1$, $\mathbf{E}[C]$ is bounded by $\frac{c^3}{1-c}$, which does not depend on n . Then the number of cycles is bounded.

4 The emergence of the giant component

The most spectacular result, when dealing with phase transitions in Erdős-Rényi graphs concerns the study of the size of the largest component, and we can exhibit several behaviors very precisely, but we will only study the coarse behavior, when $p = c/n$, where c is a constant.

In this paragraph, for a vertex u , we denote by $C(u)$ the connected component to which u belongs, and C_1 is the largest connected component, C_2 the second largest component.

Theorem 12. *Depending on c , the following cases can occur:*

(i) (sub-critical regime) *If $c < 1$, then there exists a depending on c such that*

$$\lim_{n \rightarrow \infty} \mathbf{P}(|C_1| \leq a \ln n) = 1.$$

(ii) (critical regime) *If $c = 1$, then there is a constant $\kappa > 0$ such that for all $a > 0$,*

$$\lim_{n \rightarrow \infty} \mathbf{P}(|C_1| \geq an^{2/3}) \leq \frac{\kappa}{a^2}.$$

(iii) (super-critical regime) *If $c > 1$, let p_e be the unique positive solution of $x = e^{-c(1-x)}$. There exists a constant a' depending of c such that for all $\delta > 0$,*

$$\lim_{n \rightarrow \infty} \mathbf{P}\left(\left|\frac{|C_1|}{n} - (1 - p_e)\right| \leq \delta \text{ and } |C_2| \leq a' \ln n\right) = 1.$$

4.1 Analysis of one connected component and branching processes

The key tool to study connected components is the branching process. Indeed, if one studies the size of the connected component $C(u)$ containing vertex u , then, one can mimic the behavior of the BFS (breadth-first-search) algorithm.

Vertices can be live (queued vertices), neutral or dead (popped vertices).

- Initially (at time $t = 0$), every vertex is neutral except u , who is live.
- At each time t , we take one live vertex w in the queue, pop it and queue all its neighbors that are still neutral. Then those vertices become live and w becomes dead.
- The procedure ends when the queue is empty, and the dead vertices correspond to $C(u)$.

Let us denote by $L(t)$, $N(t)$ and $D(t)$ respectively the number of live, neutral and dead nodes at time t . Let $Z(t)$ be the number of vertices added in the queue at time t and T be the first time when there is no live vertex.

We have $L(0) = 1$, $D(0) = 0$ and $N(0) = n - 1$; moreover, we have the following recursion:

$$L(t) = L(t-1) - 1 + Z(t), \quad N(t) = N(t-1) - Z(t) \text{ and } D(t) = t.$$

In other words, $N(t) = n - t - L(t)$ and $Z(t)$ is found by checking the adjacency between one vertex and $N(t)$ vertices, that is

$$Z(t) \sim \mathcal{B}\text{in}(N(t-1), p) = \mathcal{B}\text{in}(n - t + 1 - L(t-1), p).$$

Comparison of the graph process with the binomial process A binomial process is when $Z \sim \mathcal{B}\text{in}(n, p)$. Here, contrary to the graph process, n does not change with the size of the branching.

We denote by $T_{n,p}^{\text{bin}}$ the size of the binomial branching process with parameters n and p , and $T_{n,p}^{\text{gr}}$ the graph process for a vertex in $\mathcal{G}(n, p)$.

We have the following results:

Lemma 6. For any $k \in \mathbb{N}$,

$$\mathbf{P}(T_{n-k,p}^{bin} \geq k) \leq \mathbf{P}(T_{n,p}^{gr} \geq k) \leq \mathbf{P}(T_{n-1,p}^{bin} \geq k) \leq \mathbf{P}(T_{n,p}^{bin} \geq k).$$

Proof: The last inequality is obvious as $\mathcal{B}\text{in}(n, p) \geq \mathcal{B}\text{in}(n-1, p)$. \square

4.2 The sub-critical regime

Let $p = c/n$ with $c < 1$.

$$\begin{aligned} \mathbf{P}(T_{n,p}^{gr} \geq u) &\leq \mathbf{P}(T_{n,p}^{bin} \geq u) \\ &\leq \mathbf{P}(\mathcal{B}\text{in}(nu, p) \geq u-1) \\ &\leq \mathbf{P}(\mathcal{B}\text{in}(nu, p) \geq uc(1 + \frac{u(1-c)-1}{uc})) \\ &\leq e^{-\frac{uc}{3} \left(\frac{u(1-c)-1}{uc} \right)^2} \quad \text{Chernoff bound} \\ &\leq e^{\frac{2(1-c)}{3c}} e^{-\frac{u(1-c)^2}{3c}} \end{aligned}$$

Set $u = a \ln n$, then

$$\mathbf{P}(T_{n,p}^{gr} \geq a \ln n) \leq e^{\frac{2(1-c)}{3c}} n^{-a \frac{(1-c)^2}{3c}}.$$

Now, if we choose $a = \frac{4c}{(1-c)^2}$, then we have $\mathbf{P}(|C(u)| \geq a \ln n) \leq e^{\frac{2(1-c)}{3c}} n^{-4/3}$ and

$$\mathbf{P}(C_1 \geq a \ln n) \leq \sum_u \mathbf{P}(|C(u)| \geq a \ln n) \leq e^{\frac{2(1-c)}{3c}} n^{-1/3} \xrightarrow{n \rightarrow \infty} 0.$$

4.3 The super-critical regime: $c > 1$

Let $k^- = a' \ln n$ and $k^+ = n^{2/3}$.

There are small and giant components only

Lemma 7. For each vertex v , with high probability,

- either the branching process from v ends before k^- steps (i.e. $|C(v)| \leq k^-$);
- or $\forall k, k^- \leq k \leq k^+$, there are at least $\frac{(c-1)k}{2}$ live vertices ($L_v(k) \geq \frac{(c-1)k}{2}$).

We call a *bad node* a node that does not satisfy one of those two properties.

Proof: Let v be a vertex. Either the branching process from v ends in less than k^- steps, or in more than k^- steps. Vertex v can only be a bad node in the second case, and it is a bad node if there exists $k \in]k^-, k^+]$ such that $L_v(k) < \frac{(c-1)k}{2}$. This means that the number of visited nodes at step k is less than $k + \frac{(c-1)k}{2} = \frac{(c+1)k}{2}$.

Let $B(v, k)$ be the event “ v is a bad node at step k ” (there are less than $\frac{(c+1)k}{2}$ visited nodes).

$$\begin{aligned} \mathbf{P}(B(v, k)) &\leq \mathbf{P}\left(\sum_{i=1}^k \text{Bin}\left(n - \frac{(c+1)k}{2}, \frac{c}{n}\right) \leq \frac{(c+1)k}{2} - 1\right) \\ &\leq \mathbf{P}\left(\text{Bin}\left(k\left(n - \frac{(c+1)k}{2}\right), \frac{c}{n}\right) \leq \frac{(c+1)k}{2} - 1\right) \\ &\leq \mathbf{P}\left(\text{Bin}\left(k\left(n - \frac{(c+1)k^+}{2}\right), \frac{c}{n}\right) \leq \frac{(c+1)k}{2}\right) \end{aligned}$$

We now use a Chernoff bound : $\mathbf{E}[\text{Bin}(k(n - \frac{(c+1)k^+}{2}), \frac{c}{n})] = ck(1 - \frac{(c+1)k^+}{2n})$ and we choose δ such that $(1 - \delta)ck(1 - \frac{(c+1)k^+}{2n}) = \frac{(c+1)k}{2}$:

$$\delta = 1 - \frac{(c+1)}{c(2 - (c+1)k^+/n)} \xrightarrow{n \rightarrow \infty} 1 - \frac{(c+1)}{2c} = \frac{c-1}{2c}.$$

So,

$$\mathbf{P}(B(v, k)) \xrightarrow{n \rightarrow \infty} p \leq \exp\left(-\frac{(c-1)^2}{8c}k\right)$$

and more precisely, after computations,

$$\mathbf{P}(B(v, k)) \leq \exp\left(-\left(\frac{(c-1)^2}{8c} + O(n^{-1/3})\right)k\right)$$

As a consequence, the probability that v is a bad node is bounded by

$$\begin{aligned} \mathbf{P}(\cup_{k=k^-}^{k^+} B(v, k)) &\leq \sum_{k=k^-}^{k^+} e^{-\left(\frac{(c-1)^2}{8c} + O(n^{-1/3})\right)k} \\ &\leq n^{2/3} e^{-\left(\frac{(c-1)^2}{8c} + O(n^{-1/3})\right)k^-} \\ &\leq n^{2/3} e^{-\left(\frac{(c-1)^2}{8c} + O(n^{-1/3})\right)a' \ln n} = n^{2/3} n^{-\frac{(c-1)^2}{8c}a' + O(n^{-1/3})}. \end{aligned}$$

With $a' = \frac{16c}{(c-1)^2}$, this probability is less than $n^{-4/3}$ and the probability that there exists a bad node is then less than $n^{-1/3}$. \square

As a consequence, the probability that there exists a bad node in the graph tends to 0 with the size of the graph.

We call a *small vertex* a vertex satisfying the first property and a *large vertex* a vertex satisfying the second.

There is at most one giant component Suppose that u and v are two large vertices. Let $U(u)$ and $U(v)$ be the sets of live vertices after k^+ steps of the branching processes from u and from v . We know that $U(u) \cap U(v) = \emptyset$. Moreover, $|U(u)| \geq \frac{(c-1)k^+}{2}$ and $|U(v)| \geq \frac{(c-1)k^+}{2}$.

$$\begin{aligned} \mathbf{P}(C(u) \neq C(v)) &\leq \mathbf{P}(\text{there is no arc between } U(u) \text{ and } U(v)) \\ &\leq (1-p)^{|U(u)| \cdot |U(v)|} \\ &\leq (1-p)^{\left(\frac{(c-1)k^+}{2}\right)^2} \\ &\leq e^{-p\left(\frac{(c-1)k^+}{2}\right)^2} \quad [(1-p)^x \leq e^{-px}] \\ &\leq e^{-\frac{(c-1)^2 c}{4} n^{1/3}} = o(n^{-2}) \end{aligned}$$

Consequently, $\mathbf{P}(\text{there are several giant components}) = o(1)$.

There is one giant component, of size $(1 - p_e)n$ To prove this result, we first compute the number of small vertices, and use the convergence of the binomial branching process to a Poisson branching process.

Let N_s the number of small vertices.

Let T^- be the size of a Galton-Watson branching process with offspring law $\mathcal{Bin}(n - k^-, c/n)$, T^+ be the size of a Galton-Watson branching process with offspring law $\mathcal{Bin}(n, c/n)$, and $T = |C(u)|$. We have

$$\mathbf{P}(T^+ \leq k^-) \leq \mathbf{P}(T \leq k^-) \leq \mathbf{P}(T^- \leq k^-).$$

But, when k is fixed, $\mathbf{P}(T^- \leq k^-) \xrightarrow{n \rightarrow \infty} \mathbf{P}(T_c^{poi} \leq k)$. The same holds for T^+ : $\mathbf{P}(T^+ \leq k^-) \xrightarrow{n \rightarrow \infty} \mathbf{P}(T_c^{poi} \leq k)$. Now, when k grows to ∞ , $\mathbf{P}(T_c^{poi} \leq k) \xrightarrow{k \rightarrow \infty} p_e$. Then

$$\mathbf{E}[N_s] = (p_e + o(1))n.$$

This is not enough to finish the proof: we need to prove that $\mathbf{P}(|N_s - \mathbf{E}[N_s]| \geq \delta \mathbf{E}[N_s]) \leq \epsilon$. We will use the Tchebychev inequality.

Let S_u the random variable that is equal to 1 if u is a small vertex and to 0 otherwise. We have $N_s = \sum_u S_u$ and

$$\begin{aligned} \mathbf{E}[N_s^2] &= \sum_u \mathbf{E}[S_u] + \sum_{u \neq v} \mathbf{E}[S_u S_v] \\ &= \mathbf{E}[N_s] + \sum_v \mathbf{P}(S_v = 1) \sum_{u \neq v} \mathbf{P}(S_u = 1 | S_v = 1). \end{aligned}$$

But for each v ,

$$\begin{aligned} \sum_{u \neq v} \mathbf{P}(S_u = 1 | S_v = 1) &= \sum_{u \neq v, u \in C(v)} \mathbf{P}(S_u = 1 | S_v = 1) + \sum_{u \notin C(v)} \mathbf{P}(S_u = 1 | S_v = 1) \\ &\leq k^- + (p_e + o(1))n = (p_e + o(1))n. \end{aligned}$$

Then

$$\begin{aligned} \text{Var}(N_s) &\leq \mathbf{E}[N_s] + n^2(p_e + o(1))^2 - \mathbf{E}[N_s]^2 \\ &\leq \mathbf{E}[N_s] + o(\mathbf{E}[N_s]^2) \end{aligned}$$

And

$$\mathbf{P}(|N_s - \mathbf{E}[N_s]| \geq \delta \mathbf{E}[N_s]) \leq \frac{\text{Var}(N_s)}{\delta^2 \mathbf{E}[N_s]^2} = \frac{1}{\delta^2} \left(\frac{1}{\mathbf{E}[N_s]} + o(1) \right) = o(1).$$

4.4 Application: epidemic models

Random graphs can be seen as a model for epidemic processes. Consider the Reed-Frost model: consider a population of n individuals.

- At time 0, a single individual is infected.
- when an individual is infected, it is infectious during one time step, and after, it is removed (dead or immunized...).

- While infectious, it can infect every other healthy individual with probability p , and independently of the other infections.

More formally, let $Z_u(t) \in \{S, I, R\}$ be the state (susceptible, infected, removed) of vertex u at time t , and $Z(t) = (Z_u(t))_u$ the global state at time t . We denote by $S(z)$, $I(z)$ and $R(z)$ the number of susceptible, infected, removed vertices in state z .

The process can be modeled by a Markov chain:

$$\mathbf{P}(Z(t+1) = z' \mid Z(t) = z) = \binom{S(z)}{I(z')} (1-p)^{I(z)S(z')} [1 - (1-p)^{I(z)}]^{I(z')}$$

if $z_i \in \{I, R\} \Rightarrow z'_i = R$ and $z_i = S \Rightarrow z'_i \in \{S, I\}$.

This model can also be studied using Erdős-Rényi graphs: if u is originally infected, then the size of the epidemic is the size of the connected component $C(u)$ in $\mathcal{G}(n, p)$. If $p < 1/n$, then a small part of the individual will be infected. If $p > 1/n$, then with probability p_e , a small part of the population will be infected, and with probability p_e , a large part of the population will be infected.

5 Sequence of the degrees

The average degree in $G_{n,p}$ is $\bar{d} = p(n-1)$. In this paragraph, we show that the distribution of the degrees is, with high probability, distributed around this average.

Using the Chernoff bounds for vertex i , where d_i is the degree of vertex i , we have

$$\begin{aligned} \mathbf{P}(|d_i - \bar{d}| \geq \epsilon \bar{d}) &\leq 2e^{-\frac{\bar{d}\epsilon^2}{3}} \\ &\leq 2e^{-\frac{4 \ln n}{3}} = 2n^{-4/3} \end{aligned}$$

where $\epsilon = 2\sqrt{\frac{\ln n}{d}}$.

Then,

$$\mathbf{P}(\max_{i=1}^n |d_i - \bar{d}| \geq \epsilon \bar{d}) \leq 2n^{-1/3}.$$

This behavior is not representative of many examples of graphs. For example, we would like the distributions of the degrees to follow a *power law*, that is, the number of vertices of degree i is roughly proportional to $i^{-\beta}$ for some $\beta > 2$.

6 Small world graphs

Exercise 8

Routing in small-world graphs

In 1967, The sociologist Stanley Milgram published the results of one of its experiments. He asked several people to transmit an envelope to another person, only knowing the following information about the recipient: his profession, name and address. The envelope had to be transmitted only by relationship relations and could not be sent directly. Most of the envelopes reached their destination, and in most of the cases, the number of intermediates between the sender and the recipient was at most 6. This is what is called the *small world phenomenon*.

In Erdős-Rényi graphs, we say that the small world appears when the diameter of the graph is logarithmic in the number of vertices.

Here, we are interested in the routing in some kind of graphs with a small diameter (which we assume in this exercise), and particularly in the Kleinberg model. The vertices of the graph are on a grid $\{1, \dots, m\} \times \{1, \dots, m\}$. Two vertices are adjacent if $|u, v| = 1$ (where $|\cdot|$ is the L_1 distance). We add to this grid one (directed) *shortcuts* by vertex. A vertex u has a shortcut toward v with probability $|u - v|^{-\alpha} / \sum_{w \neq u} |u - w|^{-\alpha}$.

Consider the following greedy routing: at each step, go to the neighbor that is the nearest from the destination. Fix a source u and a destination v . Set $u(t)$ the vertex that is reached after t steps. We denote $T_{alg}(u, v)$ the number of steps to reach v from u with this algorithm.

The case where $\alpha = 2$ is very interesting: the greedy routing is efficient. We say that $u(t)$ is in phase j step t if $2^j < |u(t) - v| \leq 2^{j+1}$.

1. What is the probability that $u(t+1)$ belongs to a better phase than $u(t)$? Show that this probability is greater than $1/72(1 + \log(2m))$.
2. Deduce that $\mathbf{E}(T_{alg}(u, v)) = O(\log(n)^2)$.

We now suppose that $\alpha \neq 2$ and we show that the greedy routing is not efficient anymore.

3. Show that when $\alpha < 2$, $\mathbf{E}(T_{alg}(u, v)) = \Omega(m^{\frac{2-\alpha}{3}})$. One can consider the last shortcut taken by a routing of length t and the distance of this shortcut to v .
4. Show that when $\alpha > 2$, the routing algorithm from u to v terminates in average with $\mathbf{E}(T_{alg}(u, v)) = \Omega(|u - v|^\gamma)$ steps, where $\gamma = (\alpha - 2)/(\alpha - 1)$. One can first compute the probability to have a shortcut of length less than d and bound the probability to have a routing on at most t steps between two vertices at distance $td + 1$.

Exercise 9

Graphs with large girth and chromatic number

Random graphs can also be used to show the existence of some graph satisfying some property. For example, given k and ℓ , one wishes to show that there exists a graph with girth at least k and chromatic number at least ℓ . We recall that the girth $g(G)$ of a graph G is the smallest length of a cycle in and the chromatic number $\chi(G)$ the smallest number k such that the graph is k -colorable.

Set $\epsilon < \frac{1}{\ell}$ and $p = n^{\epsilon-1}$ and consider a random graph $G_{n,p}$ in $\mathcal{G}(n, p)$.

1. What is the expectation of X , the number of cycles of length at most ℓ in $G_{n,p}$?
2. Show that there exists n such that $\mathbf{P}(X \geq n/2) < 1/2$.

We now bound the size of an independent set. Let $\alpha(G)$ be the size of the largest independent set of G .

3. If in a graph G there is no independent set of size larger than $a > 2$, what is a lower bound on the chromatic number?
4. Let $a > 2$. Give an upper bound of $\mathbf{P}(\alpha(G_{n,p}) \geq a)$. Show that with $a = \lceil \frac{3 \ln n}{p} \rceil$, it is possible to choose n such that $\mathbf{P}(\alpha(G_{n,p}) \geq a) < 1/2$.

5. Deduce that there exists a graph G with $\alpha(G) < a$ and at most $n/2$ cycles of length at most ℓ .
6. Construct a graph G^* from G (by removing some vertices) such that $\alpha(G) < a$ and $g(G^*) > \ell$. Conclude.

7 Other models of random graphs

7.1 The configuration model

In this model, the sequence of the degrees is fixed so one can choose any distribution for the degrees.

7.1.1 Probability space

Let $n \in \mathbb{N}$ and $\mathbf{d} = (d_1, \dots, d_n) \in \mathbb{N}^n$ be a sequence of integers with an even sum ($\sum_{i=1}^n d_i = 2m$). Then $\mathcal{G}^*(n, \mathbf{d})$ is a probability space on the configurations obtained by pairing the $2m$ elements. This corresponds to the fact that each vertex has d_i *semi-edges* that are numbered from 1 to d_i .

A *configuration* is a pairing of the semi-edges leads to a *multigraph* (with self-loops) by forgetting the numbering of the semi-edges.

- $\mathcal{G}(n, \mathbf{d})$ is the space of the simple graphs obtained from the configurations, with a uniform distribution.
- $\mathcal{G}^*(n, \mathbf{d})$ is the space of multigraphs when the configurations are uniformly distributed.

Exercise 10

What are the configurations and possible multigraphs for $n = 3$ and $\mathbf{d} = (2, 2, 2)$? Is it possible to construct a simple graph?

We now explain how to generate a simple graph according to the uniform distribution, and that asymptotically, there exists a simple graph for a given degree sequence.

7.1.2 Construction

Let us first construct a graph in $\mathcal{G}^*(n, \mathbf{d})$ with Algorithm 1, that we will denote $G^*(n, \mathbf{d})$.

Algorithm 1: Construction of a graph in $\mathcal{G}^*(n, \mathbf{d})$.

```

begin
   $k \leftarrow 0$   while  $k < m$  do
    Choose uniformly at random a semi-edge in the  $2m - 2k$  remaining edges;
    Choose uniformly at random a semi-edge in the  $2m - 2k - 1$  remaining edges;
    Form an edge with those two semi-edges;
     $k \leftarrow k + 1$ 
  end
end

```

Then simple graphs in $\mathcal{G}(n, \mathbf{d})$, denoted $G(n, \mathbf{d})$, can be constructed with Algorithm 2.

We have to check that this algorithm is correct, that is that it generates a graph uniformly at random and that it will terminate in finite time, that is, that there exists indeed a positive proportion of simple graphs from the configuration.

Let us first focus on the correction. Let G^* be a multigraph obtained from a configuration.

Algorithm 2: Construction of a graph in $\mathcal{G}(n, \mathbf{d})$.

```

begin
  repeat
    Generate a graph in  $\mathcal{G}^*(n, \mathbf{d})$ ;
  until this graph is simple ;
end

```

Number of configurations: all the semi-edges have different names, so, according to the procedure of construction of a configuration, the number of configurations is

$$\frac{2m \times (2m - 1)}{2} \times \frac{(2m - 2) \times (2m - 3)}{2} \times \dots \times \frac{2 \times 1}{2} \times \frac{1}{m!} = \frac{(2m)!}{2^m m!},$$

where the $m!$ is for the number of orders to choose the edges.

Number of configurations leading to G^* Let us denote by m_{ij} the number of multiple edges between vertices i and j (the edge is simple if $m_{ij} = 1$) and $m_{ii}/2$ the number of self-loops around i . The number of configuration leading to G^* is then

$$\prod_{i=1}^n \frac{d_i! \left(\frac{m_{ii}}{2}\right)!}{(\prod_{j=1}^n m_{i,j}!) 2^{m_{ii}/2}}.$$

Indeed, consider for each vertex the number of possibilities to join m_{ij} times with j , for each of the $d_i!$ orders of the semi-edges of i . If $j \neq i$, then there are $m_{ij}!$ different possibilities. If $j = i$, we count the possibilities for the self-loops. There are $m_{ii} - 1$ possibilities for the first semi-edge, then $m_{ii} - 3$ for the second edge (the first semi-edge semi-edge is chosen arbitrarily for each pair) and so on. Finally, there are $(m_{ii} - 1) \times (m_{ii} - 3) \times \dots \times 1 = \frac{m_{ii}! 2^{m_{ii}/2}}{\left(\frac{m_{ii}}{2}\right)!}$, hence the result.

Therefore,

$$\mathbf{P}(G^*(n, \mathbf{d}) = G^*) = \frac{m! 2^m \prod_{i=1}^n (d_i! \left(\frac{m_{ii}}{2}\right)!)}{(2m)! \prod_{i=1}^n (\prod_{j=1}^n m_{i,j}!) 2^{m_{ii}/2}}.$$

We can check this result on the previous example.

If G^* is a simple graph, the formula becomes ($m_{ii} = 0$ and $m_{ij} = 1$ for $j \neq i$)

$$\mathbf{P}(G^*(n, \mathbf{d}) = G^* \text{ simple}) = \frac{m! 2^m \prod_{i=1}^n d_i!}{(2m)!},$$

and this probability is the same for every simple graph.

To conclude, if G is a simple graph, the probability that the second algorithm outputs G is

$$\begin{aligned} \mathbf{P}(G(n, \mathbf{d}) = G) &= \mathbf{P}(G^*(n, \mathbf{d}) = G \mid G \text{ is simple}) \\ &= \frac{m! 2^m \prod_{i=1}^n d_i!}{(2m)!} \frac{1}{\mathbf{P}(G \text{ is simple in } \mathcal{G}^*(n, \mathbf{d}))}. \end{aligned}$$

So the distribution is the uniform distribution.

We now focus on the second problem by counting the number of self-loops and cycles of length 2. To keep the computation quite simple, we consider the case of regular graphs. The results holds under more general assumptions.

7.1.3 Number of small cycles

We assume that $\mathbf{d} = (r, r, r, \dots, r)$. Then $m = \frac{rn}{2}$.

Given a multigraph G , define $Z_k(G)$ the number of cycles of length k in G . Then

- $Z_1(G)$ is the number of self-loops of G ;
- $Z_2(G)$ is the number of parallel pairs in G .

The graph is simple if $Z_1(G) = Z_2(G) = 0$.

Fix k edges in a multigraph that form a simple graph and more precisely fix k corresponding numbered edges in the configuration space. Let W be a random configuration.

The probability that W contain those edges is

$$p_k = \frac{(rn/2)! 2^{rn/2}}{(rn)!} \times \frac{(rn - 2k)!}{(rn/2 - k)! 2^{rn/2 - k}}.$$

Let a_k be the number of potential cycles of length k . Then $\mathbf{E}(Z_k) = a_k p_k$. Each cycle can be described in $2k$ manners, starting point plus direction, and for each vertex in the cycle, there are $r(r - 1)$ possible choices for the semi-edges. Then $2ka_k = \frac{n!}{(n-k)!} [r(r - 1)]^k$.

Now, as k is fixed, we have $p_k \sim 2^k \frac{(rn)^{-2k}}{(\frac{rn}{2})^{-k}} = (rn)^{-k}$ and $a_k \sim \frac{n^k}{2k} r^k (r - 1)^k$. As a consequence,

$$\mathbf{E}[Z_k(G^*(n, \mathbf{d}))] \sim \frac{(r - 1)^k}{2k} = \lambda_k.$$

Let $x^{(k)} = x(x - 1) \cdots (x - k + 1)$ and denote by C_k the set of cycles of length k .

$$\begin{aligned} \mathbf{E}[Z_k^{(2)}(G^*(n, \mathbf{d}))] &= \sum_{c \in C_k} \sum_{c' \in C_k \setminus \{c\}} \mathbf{P}(c \text{ and } c' \text{ are cycles in } G^*(n, d)) \\ &= \sum_{c \in C_k} \mathbf{P}(c \text{ is a cycle in } G^*(n, d)) \\ &\quad \sum_{c' \in C_k \setminus \{c\}} \mathbf{P}(c' \text{ is a cycle in } G^*(n, d) \mid c \text{ is a cycle in } G^*(n, d)) \\ &= a_k p_k \left[\sum_{c' \neq c, c' \cap c \neq \emptyset} \mathbf{P}(c' \text{ is a cycle in } G^*(n, d) \mid c \text{ is a cycle in } G^*(n, d)) \right. \\ &\quad \left. + \sum_{c' \neq c, c' \cap c = \emptyset} \mathbf{P}(c' \text{ is a cycle in } G^*(n, d) \mid c \text{ is a cycle in } G^*(n, d)) \right] \end{aligned}$$

The first term of the sum is negligible before the second term when k is fixed. The second term is equivalent to $a_k p_k$ (it is asymptotically distributed as $\mathbf{E}[Z_k(G^*(n - k, d))]$). As a consequence, $\mathbf{E}[Z_k^{(2)}] \xrightarrow{n \rightarrow \infty} \lambda_k^2$.

Using the same kind of argument, we can show that $\forall \ell \in \mathbb{N}$ and $(k_i) \in \mathbb{N}^\ell$,

$$\mathbf{E}[Z_1^{(k_1)}(G^*(n - k, d)) Z_2^{(k_2)}(G^*(n - k, d)) \cdots Z_\ell^{(k_\ell)}(G^*(n - k, d))] \xrightarrow{n \rightarrow \infty} \lambda_1^{k_1} \lambda_2^{k_2} \cdots \lambda_\ell^{k_\ell}.$$

Lemma 8. *If $\forall \ell \in \mathbb{N}$, $\forall (k_i) \in \mathbb{N}^\ell$, and $\forall (j_i) \in \mathbb{N}^\ell$,*

$$\mathbf{E}[Z_{j_1}(n)^{(k_1)} Z_2^{(k_2)}(n) \cdots Z_{j_\ell}(n)^{(k_\ell)}] \xrightarrow{n \rightarrow \infty} \lambda_{i_1}^{k_1} \lambda_2^{k_2} \cdots \lambda_{i_\ell}^{k_\ell},$$

then $Z_i \xrightarrow{n \rightarrow \infty} \text{Poi}(\lambda_i)$ in distribution, and $(Z_i)_i$ is mutually independent.

Proof: We only give the sketch of the proof. First, let $X \sim \mathcal{Poi}(\lambda)$. Its generating function is $g_X(s) = e^{\lambda(s-1)}$ and

$$g_X^{(k)}(1) = \mathbf{E}[X(X-1)\cdots(X-k+1)] = \lambda^k e^{\lambda(1-1)} = \lambda^k.$$

For any variable $Z_i(n)$, we have for all $k \in \mathbb{N}$, $\mathbf{E}[Z_i^{(k)}(n)] \xrightarrow{n \rightarrow \infty} \lambda_i$. So, as the generating function characterizes the distribution, $Z_i(n) \xrightarrow{n \rightarrow \infty} X$ in distribution.

Let $X_i \sim \mathcal{Poi}(\lambda_i)$ be mutually independent random variables.

To deduce the mutual independence, it suffices to observe that

$$\mathbf{E}[Z_1^{(k_1)}(n)Z_2^{(k_2)}(n)\cdots Z_\ell^{(k_\ell)}(n)] \xrightarrow{n \rightarrow \infty} \mathbf{E}[X_1^{(k_1)}]\mathbf{E}[X_2^{(k_2)}]\cdots\mathbf{E}[X_\ell^{(k_\ell)}] = \mathbf{E}[X_1^{(k_1)}X_2^{(k_2)}\cdots X_\ell^{(k_\ell)}].$$

□

We can now prove the following theorem:

Theorem 13. *Let G be a random multigraph of the configuration model with degree sequence $\mathbf{d} = (r, \dots, r)$. Then asymptotically, $\mathbf{P}(G \text{ is simple}) = e^{-\frac{r^2-1}{4}}$.*

Proof: Asymptotically, $Z_1 \sim \mathcal{Poi}(\lambda_1) = \mathcal{Poi}(\frac{r-1}{2})$ and $Z_2 \sim \mathcal{Poi}(\lambda_2) = \mathcal{Poi}(\frac{(r-1)^2}{4})$. Then

$$\mathbf{P}(Z_1 = 0, Z_2 = 0) = \mathbf{P}(Z_1 = 0)\mathbf{P}(Z_2 = 0) = e^{-\lambda_1 - \lambda_2} = e^{-\frac{r^2-1}{4}}.$$

□

Theorem 14. *Any property that holds a.a.s in $\mathcal{G}^*(n, \mathbf{d})$ holds asymptotically in $\mathcal{G}(n, \mathbf{d})$.*

Proof: Let \mathcal{P} be a property.

$$\begin{aligned} \mathbf{P}(G(n, \mathbf{d}) \text{ does not satisfy } \mathcal{P}) &= \mathbf{P}(G^*(n, \mathbf{d}) \text{ does not satisfy } \mathcal{P} | G^*(n, \mathbf{d}) \text{ is simple}) \\ &= \frac{\mathbf{P}(G^*(n, \mathbf{d}) \text{ does not satisfy } \mathcal{P} \text{ and is simple})}{\mathbf{P}(G^*(n, \mathbf{d}) \text{ is simple})} \\ &= \frac{\mathbf{P}(G^*(n, \mathbf{d}) \text{ does not satisfy } \mathcal{P})}{\mathbf{P}(G^*(n, \mathbf{d}) \text{ is simple})} \rightarrow 0 \end{aligned}$$

□

The reverse is not true. For example it does not hold for the property “not containing a loop”.

7.1.4 Erased configuration model

A simple graph can be obtained from a multigraph by erasing all the self-loops and merging the multi-edges. The degree of vertex i then becomes

$$D_i^{(er)} = d_i - 2s_i - m_i,$$

where s_i is the number of self-loops around i and m_i the number of multi-edges merged for i . Let

- $p_k = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{d_i=k\}}$ be the proportion of vertices of degree k in the initial graph and
- $p_k^{(n)} = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{D_i^{(er)}=k\}}$ the proportion of vertices of degree k in the erased graph.

Theorem 15. For all $\epsilon > 0$, $\mathbf{P}(\sum_{k=0}^{\infty} |p_k - p_k^{(er)}| \geq \epsilon) \xrightarrow{n \rightarrow \infty} 0$.

Proof: $\sum_{k=0}^{\infty} |p_k - p_k^{(er)}| \leq \frac{1}{n} \sum_{k=0}^{\infty} \sum_{i=1}^n |\mathbf{1}_{\{d_i=k\}} - \mathbf{1}_{\{D_i^{(er)}=k\}}|$. But,

$$\begin{aligned} \mathbf{1}_{\{D_i^{(er)}=k\}} - \mathbf{1}_{\{d_i=r\}} &= \mathbf{1}_{\{D_i^{(er)}=k, d_i > k\}} - \mathbf{1}_{\{d_i < r, D_i^{(er)}=k\}} \\ &= \mathbf{1}_{\{s_i+m_i > 0\}} (\mathbf{1}_{\{D_i^{(er)}=k\}} - \mathbf{1}_{\{d_i=r\}}) \end{aligned}$$

So $|\mathbf{1}_{\{D_i^{(er)}=k\}} - \mathbf{1}_{\{d_i=r\}}| \leq \mathbf{1}_{\{s_i+m_i > 0\}} (\mathbf{1}_{\{D_i^{(er)}=k\}} + \mathbf{1}_{\{d_i=r\}})$ and

$$\begin{aligned} \sum_{k=0}^{\infty} |p_k - p_k^{(er)}| &\leq \frac{1}{n} \sum_{k=0}^{\infty} \sum_{i=1}^n |\mathbf{1}_{\{d_i=k\}} - \mathbf{1}_{\{D_i^{(er)}=k\}}| \\ &\leq \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{s_i+m_i > 0\}} \sum_{k=0}^{\infty} (\mathbf{1}_{\{D_i^{(er)}=k\}} + \mathbf{1}_{\{d_i=r\}}) \\ &\leq \frac{2}{n} \sum_{i=1}^n (s_i + m_i). \end{aligned}$$

Then

$$\begin{aligned} \mathbf{P}(\sum_{k=0}^{\infty} |p_k - p_k^{(er)}| \geq \epsilon) &\leq \mathbf{P}(2 \sum_{i=1}^n s_i + m_i \geq \epsilon n) \\ &\leq \frac{2\mathbf{E}[Z_1] + 4\mathbf{E}[Z_2]}{\epsilon n} \\ &\leq \frac{2\lambda_1 + 4\lambda_2}{\epsilon n} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

□

7.2 Preferential attachment graphs

Preferential attachment graphs are class of graphs with a power law distribution of the degrees.

A random variable X has a power law distribution with parameter β if

$$\mathbf{P}(X = i) = C i^{-\beta}$$

where C is a normalizing constant.

7.2.1 Construction of the graph

Initially, set $G(0) = (V(0), E(0))$ be a graph with vertices $V(0)$ and edges $E(0)$. At time t , we have $V(t) = V(0) + t$ and $E(t) = E(0) + t$.

At time $t + 1$, add vertex $u(t + 1)$ and attach it as follows:

- With probability α , choose a vertex uniformly at random among $V(t)$;
- With probability $1 - \alpha$, choose a vertex v with probability $\frac{d_t(v)}{2E(t)}$, where $d_t(v)$ is the degree of vertex v at time t .

Join $u(t + 1)$ with the chosen vertex.

1. What kind of graph is obtained?
2. How the model could be modify to create cycles?

7.2.2 Evolution of the number of vertices with a given degree

Let \mathcal{F}_t be the sigma-field containing all the information about the t first steps of the construction.

Let $X_i(t)$ be the number of vertices with degree i at time t .

3. Compute the probabilities $\mathbf{P}(X_i(t+1) = X_i(t) + a \mid \mathcal{F}_t)$, for $a \in \{-1, 0, 1\}$. Distinguish two cases: $i = 1$ and $i > 1$.

Define c_i as $c_1 = \frac{2}{3+\alpha}$ and $\frac{c_i}{c_{i+1}} = \frac{\alpha + \frac{1-\alpha}{2}(i-1)}{1+\alpha + \frac{1-\alpha}{2}i}$ for $i > 1$.

4. Show that $\frac{c_i}{c_{i+1}} = 1 - \frac{1}{i} \frac{3-\alpha}{1-\alpha} + O(\frac{1}{i^2})$.

Then, one admits (and can check) that $c_i \sim Ci^{-\beta}$ with $\beta = \frac{3-\alpha}{1-\alpha}$.

Fix $\epsilon > 0$ and set $\Delta_i = \mathbf{E}[X_i] - c_i t$.

5. Show that $\Delta_i(t+1) = \Delta_1(t) - c_1 + 1 - \alpha \frac{\mathbf{E}[X_1]}{N(t)} - (1-\alpha) \frac{\mathbf{E}[X_1]}{2E(t)}$, then that

$$\Delta_i(t+1) = \Delta_1(t) \left(1 - \frac{\alpha}{N(t)} - \frac{1-\alpha}{2E(t)}\right) + O(t^{-1}).$$

6. Finally show that $\mathbf{E}[X_1(t)] - c_1 t = o(t^\epsilon)$.

7. Using a similar argument, show that

$$\Delta_i(t+1) = \Delta_i(t) \left(1 - \frac{\alpha}{N(t)} - \frac{(1-\alpha)i}{2E(t)}\right) + O(\Delta_{i-1}(t)/t) + O(t^{-1}).$$

8. Finally show that $\mathbf{E}[X_i(t)] - c_i t = o(t^\epsilon)$.

In fact, we can show a much more stronger result: $\frac{X_i(t)}{t} \xrightarrow{t \rightarrow \infty} c_i$ almost surely.