Probabilistic Aspects of Computer Science
Random graphs

Anne Bouillard

November 28, 2020
Properties of random graphs in practice

- **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...
Why studying random graphs?

- Graphs with large size that are described by simple and local rules
- Possess (or not) properties (ou pas) of the previous slide with strong probability
- Good models for studying large graphs
  - Social networks
  - Propagation of epidemics...

We are interested in the behavior of those graphs when the number of vertices grows to infinity.
Different models of random graphs

1. **Erdős-Rényi graphs**: independent edges
2. **Configurations model**: sequence of degrees
3. **Preferential attachments**: incremental model
4. **Structured graphs**: grid with shortcuts.

Each edge is present with probability $p$.

Each vertex has a given degree, and the edges chosen at random.
Erdős-Rényi graphs: definition

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

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

Example

In \( G(n, p) \),
- 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 \( \binom{N}{m} p^m (1 - p)^{N-m} \).
Asymptotic behavior

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 $n$. 
First-order properties of graphs

Closed form formulas generated by

\[ F ::= \forall x F \mid \exists x F \mid F \lor F \mid F \land F \mid \neg F \mid x = y \mid I(x, y) \]

with the two axioms

\[ \forall x \neg I(x, x) \quad \text{and} \quad \forall x \forall y I(x, y) \iff I(y, x). \]

Example

The following properties are first-order:

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

The following properties are not first-order: \( G \) is connected, \( G \) is Hamiltonian, \( G \) is planar...
A 0-1 property for random graphs

**Theorem**

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

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

**Lemma**

$\forall r, s, \lim_{n \to \infty} P(G_{n,p} \text{ has } A_{r,s}) = 1$. 
A 0-1 property for random graphs

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

$$\Pr(A(x_i),(y_j),z) = p^r(1-p)^s$$
$$\Pr(\forall z \neg A(x_i),(y_j),z) \leq (1-p^r(1-p)^s)^n$$
$$\Pr(\exists(x_i),(y_j) \forall z \neg A(x_i),(y_j),z) \leq n^{r+s}(1-p^r(1-p)^s)^n$$
$$\Pr(G_{n,p} \text{ has } A_{r,s}) \leq 1 - n^{r+s}(1-p^r(1-p)^s)^n$$

Hence $\lim_{n \to \infty} \Pr(G_{n,p} \text{ has } A_{r,s}) = 1$. 
Building a model

We use the following results from completeness theory

- If a system has a model, then it has a denumerable model.
- 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 $G_{n+1}$. 
Equivalence

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.

**Lemma**

$G$ and $G'$ are isomorphic.

The set of vertices of $G$ and $G'$ is $\mathbb{N}$. We build an isomorphism by induction.
Construction of the isomorphism

- Let \( f \) be this isomorphism and initially set \( f(0) = 0 \).
- Let \( V = \{0, \ldots, i - 1, f^{-1}(0), \ldots, f^{-1}(i - 1)\} \) be the set of vertices where \( f \) is already defined.
  - We now define \( f(i) \) and \( f^{-1}(i) \).
- 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, \((i, j)\) is an edge in \( G \) \(\iff\) \((f(i), f(j))\) is an edge in \( G' \), and the two graphs are isomorphic.
A complete system

Lemma

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.
- 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, and 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 \( S \) of \( A_{r,s} \). Then,

\[
P(\neg A \text{ in } G_{n,p}) \leq \sum_{(r,s) \in S} P(\neg A_{r,s} \text{ in } G_{n,p}) \xrightarrow{n \to \infty} 0.
\]

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

A threshold function for the property $A$ is a function $g(n)$ such that

(i) if $\lim_{n \to \infty} \frac{p(n)}{g(n)} = 0$ (or $p \ll g$), then $\lim_{n \to \infty} P(G_{n,p(n)} \text{ has } A) = 0$.

(ii) if $\lim_{n \to \infty} \frac{g(n)}{p(n)} = 0$ (or $p \gg g$), then $\lim_{n \to \infty} P(G_{n,p(n)} \text{ has } A) = 1$.

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 $P(G_{n,p} \text{ has } A) = 0$; and if $g(n) \ll p$, then $P(G_{n,p} \text{ has } A) = 1$. 
Threshold functions: examples

<table>
<thead>
<tr>
<th>property</th>
<th>threshold function $g(n)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>contains a path of length $k$</td>
<td>$n^{-\frac{k+1}{k}}$</td>
</tr>
<tr>
<td>is not planar</td>
<td>$\frac{1}{\ln n}$</td>
</tr>
<tr>
<td>contains an Hamiltonian path</td>
<td>$\frac{n}{\ln n}$</td>
</tr>
<tr>
<td>is connected</td>
<td>$\frac{n}{\ln n}$</td>
</tr>
<tr>
<td>contains a clique of size $k$</td>
<td>$n^{-\frac{k}{k-1}}$</td>
</tr>
</tbody>
</table>
Clique of size 4

**Theorem**

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

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

$X_i = 1 \iff C_i$ is a clique of size 4.

Let $X = \sum_i X_i$. 

Cliques of size 4

- $\mathbb{E}[X] = \sum \mathbb{E}[X_i] = \left(\frac{n}{4}\right)p(n)^6 = \left(\frac{1}{24}n^4 + o(n^4)\right)p(n)^6$;
- $\mathbb{E}[X^2] = \sum \mathbb{E}[X_i] + \sum_{i \neq j} \mathbb{E}[X_i X_j]$.

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

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

$\text{Var}[X] = \left(\frac{1}{24}n^4 + o(n^4)\right)p(n)^6 + \left(o(n^8)\right)p(n)^{12} + \left(\frac{6}{24^2}n^6\right)p(n)^{11} + \left(\frac{4}{24}n^5\right)p(n)^9$. 
Clique of size 4: First and second moment method

- \( \mathbb{E}[X] = \left( \frac{1}{24} n^4 + o(n^4) \right) p(n)^6; \)
- \( \text{Var}(X) = \left( \frac{1}{24} n^4 + o(n^4) \right) p(n)^6 + (o(n^8)) p(n)^{12} + \left( \frac{6}{24^2} n^6 \right) p(n)^{11} + \left( \frac{4}{24^3} n^5 \right) p(n)^9. \)

- if \( p(n) = o(n^{-2/3}) \), then by the Markov inequality,
  \[
  P(X \neq 0) \leq \mathbb{E}[X] = \left( \frac{1}{24} n^4 + o(n^4) \right) p(n)^6 = o(1).
  \]

- if \( n^{-2/3} = o(p(n)) \), \( n^4 p(n)^6 \xrightarrow{n \to \infty} \infty \) then by the second moment method,
  \[
  P(X = 0) \leq \frac{\text{Var}(X)}{\mathbb{E}[X]^2} = O(n^{-4} p(n)^{-6}) + o(1) + O(n^{-2} p(n)^{-1}) + O(n^{-3} p(n)^{-3}) = o(1).
  \]
Threshold functions for monotone increasing properties

**Definition**

A property $\mathcal{A}$ is monotone increasing if

$$G \subseteq G' \text{ and } G \text{ satisfies } \mathcal{A} \Rightarrow G' \text{ satisfies } \mathcal{A}.$$ 

**Lemma**

If $\mathcal{A}$ is a monotone increasing property, then

$$p \leq p' \Rightarrow P(G_{n,p} \text{ satisfies } \mathcal{A}) \leq P(G_{n,p'} \text{ satisfies } \mathcal{A})$$

We use a coupling argument, and the previous construction of random graphs:

1. draw $p_e \sim \text{Unif}([0, 1])$ i.i.d for edges $e \in E$
2. we obtain $G_{n,p}$ (resp. $G_{n,p'}$) where the $e$ is an edge if $p_i \leq p$ (resp. $p_i \leq p'$).
3. $G_{n,p} \subseteq G_{n,p'}$
4. $G_{n,p}$ satisfies $\mathcal{A} \Rightarrow G_{n,p'}$ satisfies $\mathcal{A}$
Threshold functions for monotone increasing properties

**Theorem**

If $\mathcal{A}$ is monotonic increasing, then there exists a threshold function for this property.

1. Find a candidate for the threshold function: $g(n)$ such that

   $$\Pr(G_{n,g(n)} \text{ satisfies } \mathcal{A}) = \frac{1}{2}.$$ 

   This function is well-defined:
   - $\Pr(G_{n,p} \text{ satisfies } \mathcal{A})$ increases with $p$ (lemma)
   - $\Pr(G_{n,p} \text{ satisfies } \mathcal{A}) = \sum_{G \text{ satisfies } \mathcal{A}} p^{|E(G)|} (1 - p)^{|N - E(G)|}$ is continuous + intermediate value theorem

2. $k$ copies of $G(n, p)$
   - Let $G_{i,n,p}$ be $k$ independent copies of $G_{n,p}$
   - $G = \bigcup_{i=1}^{k} G_{i,n,p}$ the union of these graphs (edge-wise)
   - $G \sim G_{n,q}$ with $q = 1 - (1 - p)^k \leq kp$, so

   $$\Pr(G_{n, kp} \notin \mathcal{A}) \leq \Pr(G_{n, q} \notin \mathcal{A}) \leq \left(\Pr(G_{n, p} \notin \mathcal{A})\right)^k.$$
Threshold functions for monotone increasing properties

\[ P(G_{n,kp} \notin A) \leq P(G_{n,q} \notin A) \leq (P(G_{n,p} \notin A))^k. \]

1. \( k = \omega(n) \to \infty \) and \( p = g(n) \):

\[ P(G_{n,\omega(n)g(n)} \notin A) \leq (P(G_{n,g(n)} \notin A))^{\omega(n)} = \left(\frac{1}{2}\right)^{\omega(n)} \to 0. \]

2. \( p = g(n)/\omega(n) \):

\[ \frac{1}{2} = (P(G_{n,g(n)} \notin A)) \leq (P(G_{n,p(n)} \notin A))^{\omega(n)}, \]

so

\[ (P(G_{n,p(n)} \notin A))^{\omega(n)} \geq \left(\frac{1}{2}\right)^{1/\omega(n)} \to 1. \]
Moment generating functions

Definition

Let $X$ be a random variable on $\mathbb{N}$. Its (moment) generating functions is

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

- $g_X$ is $C^\infty$ on $] -1, 1[$
- $g_X(0) = \mathbb{P}(X = 0)$, $g_X(1) = 1$
- $\mathbb{P}(X = n) = g_X^{(n)}(0)/n!$
- $\mathbb{E}[X] = g_X'(1)$

Proposition

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 $g_{X+Y} = g_X g_Y$. 
Examples of moment generating functions

Example

- $X \sim \text{Ber}(p): g_X(s) = 1 - p + ps$;
- $X \sim \text{Bin}(n, p): g_X(s) = (1 - p + ps)^n$;
- $X \sim \text{Poi}(\lambda): g_X(s) = g_X(s) = e^{\lambda(s-1)}$;

Proposition

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.
Wald's equality: preliminary lemma

**Theorem**

Let $T$ be a non-negative integer r.v. and $(Z_i)_{i \in \mathbb{N}}$ be a sequence of 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 generating functions of $Z_1$, $T$ and $X$.

$$g_X = g_T \circ g_Z.$$

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

$$\mathbb{E}(s^{Z_1+\ldots+Z_T}) = \sum_{n=0}^{\infty} \mathbb{E}[1\{T=n\} s^{Z_1+\ldots+Z_n}] \text{ (linearity)}$$

$$= \sum_{n=0}^{\infty} \mathbb{E}[1\{T=n\}] \mathbb{E}[s^{Z_1+\ldots+Z_n}] \text{ (independence of } T \text{ and } Z_i)$$

$$= \sum_{n=0}^{\infty} \mathbb{P}(T = n)[g_Z(s)]^n \text{ (independence of the } Z_i)$$

$$= \mathbb{E}[g_Z(s)^T] = g_T(g_Z(s)).$$
Wald’s equality

**Theorem (Wald’s equality)**

Let $T$ be a non-negative integer random variable and $(Z_i)_{i \in \mathbb{N}}$ be a sequence of 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

$$E[X] = E[Z]E[T].$$

$$E[X] = g_X'(1) = g_Z'(1)g_T'(g_Z(1)) = g_Z'(1)g_T'(1) = E[Z]E[T].$$
The idea of the Chernoff bounds is to apply Markov inequality to the generating function.

**Theorem**

- \( \forall s > 1, \ P(X \geq a) \leq \inf_{s > 1} \frac{E(s^X)}{s^a} \)
- \( \forall s < 1, \ P(X \leq a) \leq \inf_{s < 1} \frac{E(s^X)}{s^a} \)

\[
\forall s > 1, \ P(X \geq a) = P(s^X \geq s^a) \leq \frac{E(s^X)}{s^a}
\]

\[
\forall s < 1, \ P(X \leq a) = P(s^X \geq s^a) \leq \frac{E(s^X)}{s^a}
\]
Special case: sum of independent Bernoulli variables

**Theorem**

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

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

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

- $g_i$: generating function of $X_i$, so $g_i(s) = 1 - p_i + p_is = 1 + p_1(s-1) \leq e^{p_i(s-1)}$.
- $g_X(s) = \prod_{i=1}^{n} g_i(s) \leq \prod_{i=1}^{n} e^{p_i(s-1)} = e^{\mu(s-1)}$.
- $\forall s > 1$, $\mathbb{P}(X \geq (1 + \delta)\mu) \leq \frac{\mathbb{E}(s^X)}{s^{(1+\delta)\mu}} \leq \frac{e^{\mu(s-1)}}{s^{(1+\delta)\mu}}$.
- with $s = 1 + \delta$, we get

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

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

A. Bouillard PACS November 28, 2020 28 / 36
**Special case: sum of independent Bernoulli variables (2)**

**Theorem**

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

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

The proof is exactly the same with $s < 1$.

- $\mathbb{P}(X \leq (1 - \delta)\mu) \leq \frac{\mathbb{E}(s^X)}{s(1-\delta)\mu} \leq \frac{e^{\mu(s-1)}}{s(1-\delta)\mu}$.
- with $s = 1 - \delta$, we get

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

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

A. Bouillard  

PACS  

November 28, 2020  29 / 36
Galton-Watson branching processes

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

- $X_0 = 1$ (root, depth 0)
- $X_n$ number of nodes at depth $n$ (population of the $n$-th generation)
- $Z_i^{(n)}$ number of children of node $i$ of the $n$-th generation. The $(Z_i^{(n)})_{i,n}$ are i.i.d.

What is the probability that the tree is finite?
**Theorem**

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

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

$Z$ is a r.v. with the same distribution as $Z_i^{(n)}$. 
Distribution of the population of the $n$-the generation

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

Generating functions:
- \( g(s) = \mathbb{E}[s^Z] \) the generating function of \( Z \)
- \( \phi_n = \mathbb{E}[s^{X_n}] \) that of \( X_n \).

Lemma

\[ \phi_{n+1} = g_Z(\phi_n). \]

- Wald equality: \( \phi_{n+1} = \phi_n \circ g_Z \).
- \( \phi_{n+1} = \phi_0 \circ g_Z \circ \cdots \circ g_Z = \phi_0 \circ g_Z^{n+1}. \)
- \( \mathbb{P}(X_0 = 1) = 1 \), so \( \phi_0(s) = s \) and \( \phi_{n+1} = g_Z^{n+1}. \)
Extinction probability as a fix-point equation

- \( p_e = \mathbb{P}(\exists n \in \mathbb{N}, \ X_n = 0) = \mathbb{P}(\bigcup_{n \in \mathbb{N}} \{X_n = 0\}) \) the extinction probability of the process
- \( \{X_n = 0\} \subseteq \{X_{n+1} = 0\} \), so \( p_e = \lim_{n \to \infty} \mathbb{P}(X_n = 0) \).

**Lemma**

\[ p_e = g_Z(p_e). \]

- We know that \( \phi_{n+1}(0) = g_Z(\phi_n(0)) \)
- we also have \( \phi_{n+1}(0) = \mathbb{P}(X_{n+1} = 0) \) and \( \phi_n(0) = \mathbb{P}(X_n = 0) \)
- Then, by continuity (\( g_Z \) is continuous on \([0, 1]\)), \( p_e = g_Z(p_e) \).
Extinction probability as a fix-point equation (2)

**Theorem (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 $\mathbb{E}[X] \leq 1$, then the equation $x = g(x)$ has a unique solution in $[0, 1]$, $x = 1$. If $\mathbb{E}[X] > 1$, then the equation $x = g(x)$ has two solutions, in $[0, 1]$, $x = 1$ and $\beta \in [0, 1]$.

- $g_Z(s) = \sum_{n \in \mathbb{N}} \mathbf{P}(Z = n)s^n$ 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 $\mathbf{P}(Z \leq 1) < 1$. 
Extinction probability as a fix-point equation (3)

- $x = 1$ is trivially a solution.
- Now, we use the convexity of $g_Z$.
- If $E[X] \leq 1$, 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 $E[X] > 1$, on an interval $[1 - \epsilon, 1[$, $g_Z(x) < x$. But $g_Z(0) \geq 0$, so there exists $\beta$ such that $\beta = g_Z(\beta)$. 

![Graphs showing the conditions for the derivative of the moment generating function $g'_Z(1)$]
Extinction probability

Theorem

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

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

- $x_n = P(X_n = 0)$
- $x_0 = 0$, so $\beta - x_0 \geq 0$
- 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 $p_e = \beta$