Random Graphs and Complex Networks

1 Branching Processes

A branching process is a simple model for something like a population evolving in time.

**Example 1.** Suppose we have some amoeba who, at each generation, each individual gives birth to a number of children given by some distribution: \( p_i = \mathbf{P} \) (individual has exactly \( i \) children). Let \( Z_n \) be the number of individuals born in the \( n^{th} \) generation. A nice way to realize the process with i.i.d offspring given by \( X \) is to set up an infinite array of i.i.d random variables \( X_{j,n} \) each distributed by the vector \( \vec{p} \) i.e. \( \mathbf{P}(X = i) = p_i \). This will represent the number of children that the individual \( j \) will have in the \( n \)-th generation. We must be careful not to count individuals who haven’t been born yet. Some thought leads us to the recursive expression for \( Z_n \):

\[
Z_n = \sum_{j=1}^{Z_{n-1}} X_{n,j}
\]

**Theorem 2.** Let \( \eta = \mathbf{P}(\exists n : Z_n = 0) \) be the probability that the population becomes extinct at some time. For a branching process with i.i.d offspring given by \( X \) as above, we will have \( \eta = 1 \) when \( \mathbf{E}(X) < 1 \) while \( \eta < 1 \) for \( \mathbf{E}(X) > 1 \). If \( \mathbf{E}(X) = 1 \) and \( \mathbf{P}(X = 1) < 1 \) then \( \eta = 1 \). Moreover, if we let \( G_X \) be the probability generating function of \( X \):

\[
G_X(s) = \mathbf{E}(s^X)
\]

Then \( \eta \) is the smallest solution in \([0,1]\) of:

\[
\eta = G_X(\eta)
\]

*Proof.** The proof uses a slick trick with probability generating functions. Let \( G_n \) be the generating function for \( Z_n \), so that \( G_n(s) = \mathbf{E}(s^{Z_n}) \). Let \( \eta_n = \mathbf{P}(Z_n = 0) = G_n(0) \) be the probability that we are extinct at time \( n \). Since once we are extinct, we stay extinct forever, we know that \( \eta_n \) is monotone non-decreasing. Moreover, \( \eta_n \to \eta \) as \( n \to \infty \) since \( \{Z_n = 0\} \uparrow \{\exists n : Z_n = 0\} \).

Now, using the fact that \( Z_n \sim \sum_{i=1}^{X} Z_{n-1} \) i.e. \( Z_n \) is distributed like \( X \) independent copies of \( Z_{n-1} \) (the \( X \) here is the number of guys in the first generation \( X = Z_1 \)), and \( Z_1 \sim X \) we get that:

\[
G_n(s) = \mathbf{E}(s^{Z_n}) = \sum_{i=0}^\infty p_i \mathbf{E}(s^{\sum_{j=1}^{Z_{n-1}} Z_{n-1}} | Z_1 = i) = \sum_{i=0}^\infty p_i \mathbf{E}(s^{\sum_{j=1}^{Z_{n-1}} Z_{n-1}} | Z_1 = i) = \sum_{i=0}^\infty p_i (G_{n-1}(s))^i = G_X(G_{n-1}(s))
\]

Plugging in \( s = 0 \) gives the recurrence relation for \( \eta_n \) through \( \eta_n = G_n(0) \):

\[
\eta_n = G_X(\eta_{n-1})
\]

Finally, since \( \eta_n \to \eta \) and since \( G_X \) is continuous in \([0,1]\) (its analytic in at least the unit disk, since the coefficients of its power series are all less than 1), by taking \( n \to \infty \) we have that \( \eta \) must satisfy:

\[
\eta = G_X(\eta)
\]

This is the main breakthrough, the rest of the result follows by a bit of careful analysis.

If \( \mathbf{P}(X = 1) = 1 \), then \( Z_n = 1 \) a.s. and so there is nothing to prove. If \( \mathbf{P}(X = 0) + \mathbf{P}(X = 1) = 1 \), but \( \mathbf{P}(X = 0) = p > 0 \), then \( G_X(s) = p + (1-p)s \), so the recurrence above gives \( \eta_n = 1 - (1-p)^n \) and \( \eta = 1 \).
Hence, with these cases covered, we may assume WLOG that $P (X \leq 1) < 1$. Let $\psi \in [0, 1]$ be any root of $\psi = G_X (\psi)$. Claim: $\eta_n \leq \psi$ for all $n$. Proof: (by induction) $\eta_0 = 0 \leq \psi$, and if $\eta_{n-1} \leq \psi$ then $\eta_n = G_X (\eta_{n-1}) \leq G_X (\psi) = \psi$ since $G_X$ is an increasing function. Hence $\eta = \lim \eta_n \leq \psi$ too. So $\eta$ must be the smallest root of $\eta = G_X (\eta)$ in $[0, 1]$.

Now, notice $G_X' = E(X (X - 1)^{s-1}) \geq 0$ so $X$ is concave up, and recall that $G_X' (1) = E(X)$. If $G_X' = E(X) < 1$, then by convexity, there can be no intersection between $G_X (s)$ and $s$ for $s \in [0, 1]$, while for $G_X' = E(X) > 1$ there must be at least one point of intersection less than 1 when we take into account the fact that $G_X (0) = P (X = 0) \geq 0$ (Draw a picture! It’s like an easy ODE’s type slope and region argument). If $G_X' (1) = E(X) = 1$ then either $G_X (0) = 0$ and $G_X (s) \equiv s$ (which is the case we already covered $X = 1$ a.s.) or $G_X (0) > 0$ and there can be no intersection except at $s = 1$ again by convexity. (Draw the picture again!)

**Theorem 3.** Let $T = \sum_{n=0}^{\infty} Z_n$ be the total size of the population. Then the generating function for $T$, $G_T (s) = E (s^T)$ satisfies:

$$G_T (s) = s G_X (G_T (s))$$

**Notice that this is not normalized if the extinction probability $\eta < 1$ (i.e. its possible that $T = \infty$), instead of the usual $G_Y (1) = 1$, we have here $G_T (1) = \eta$.**

**Proof.** This is the same type of proof as the recurrance in the above theorem, just condition on $Z_1 \sim X$ and use independence.

**Example 4.** If we consider binary branching, $X = 0$ with probability $p$, and $X = 2$ with probability $q = 1 - p$. Then:

$$G_T (s) = \frac{1 - \sqrt{1 - 4 s^2 pq}}{2sp}$$

Notice that if we plug in $s = 1$ here we get:

$$\eta = \frac{1 - |2p - 1|}{2p} = \begin{cases} 1 & p < \frac{1}{2} \\ \frac{1-p}{p} & p > \frac{1}{2} \end{cases}$$

**Remark 5.** Notice this is exactly the form of the Lagrange implicit function theorem (see notes from Combinatorics, the formula in question is $R(x) = x G(R(x)) \Rightarrow [x^n] R(x) = 0$ and $[x^n] R(x) = \frac{1}{n} [u^{n-1}] G(u)^n$ whenever $[u^0] G(u) \neq 0$) Interpreted probabilistically, this says that, under the assumption $P \{X = 0\} > 0$, we have:

$$P \{T = 0\} = 0$$
$$P \{T = n\} = \frac{1}{n} P \{X_1 + X_2 + X_3 + \ldots + X_n = n - 1\}$$

I am not sure what the combinatorial interpretation of this formula is. (Later on we will see why using the “hitting-time” theorem)

**Theorem 6.** If $E(X) = \mu$ then:

$$E(Z_n) = \mu^n$$

and moreover, $\mu^{n-1} Z_n$ is a martingale.

**Proof.** (By induction). The base case is clear, as $Z_1 \sim X$. Now suppose it holds for $n$. To see the induction step, suppose that $E(Z_{n-1}) = \mu^{n-1}$ and use the recurrence:

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}$$

We compute:

$$E(Z_n) = E \left( \sum_{i=1}^{Z_{n-1}} X_{n,i} \right)$$

$$= \sum_{k=0}^{\infty} E \left( \left( \sum_{i=1}^{Z_{n-1}} X_{n,i} \right) \left| Z_{n-1} = k \right. \right) P (Z_{n-1} = k)$$

$$= \sum_{k=0}^{\infty} k \mu P (Z_{n-1} = k)$$

$$= \mu \left( \sum_{k=0}^{\infty} k P (Z_{n-1} = k) \right)$$

$$= \mu E (Z_{n-1})$$

$$= \mu^n$$
Finally, to see that \( \mu^{-n}Z_n \) is a martingale, first notice that the recurrence \( Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i} \) shows that \( \mathbb{E}(Z_n|Z_{n-1}, \ldots, Z_1) = \mathbb{E}(Z_n|Z_{n-1}) \) and also:

\[
\mathbb{E}(Z_n|Z_{n-1}) = \mathbb{E}\left(\sum_{i=1}^{Z_{n-1}} X_{n,i}|Z_{n-1}\right) = \mu Z_{n-1}
\]

So it follows that \( \mathbb{E}(\mu^{-n}Z_n|\mu^{-(n-1)}Z_{n-1}) = \mu^{-n}\mu Z_{n-1} = \mu^{-(n-1)}Z_{n-1} \) so it is a martingale.

\( \square \)

**Remark 7.** By the martingale convergence theorem, we know that \( \mu^{-n}Z_n \rightarrow_a W \) for some random variable \( W \).

**Remark 8.** Notice that in the non-trivial case that has \( \mathbb{E}(X) = 1 \) we have that \( Z_n \rightarrow 0 \) in probability, but \( \mathbb{E}(Z_n) = 1 \) for every \( n \). This is a bit weird!

**Remark 9.** By summing, this shows \( \mathbb{E}(T) = \frac{1}{1-\mu} \).

### 1.1 Random-Walk perspective to the Branching Process

Here is an entirely different way to think about the branching process. After the branching process happens, we will count the number of people by the following exploring algorithm. We start by labeling the first dude as active. Now choose an active person. Then, in the next step, we label all his children as active and him as inactive. If there are no active people left we stop, if there are some active people left repeat. Let \( S_n \) be the number of active people after \( n \) iterations. We have:

\[
S_0 = 1 \\
S_k = S_{k-1} + X_k - 1 \\
= X_1 + \ldots + X_k - (k - 1) \\
T = \min\{t : S_t = 0\} \\
= \min\{t : X_1 + \ldots + X_k = k - 1\}
\]

Let’s find the probability distribution for this sequence. Let us create a probability distribution on \( A \cup B \) where \( A, B \) are the following sets:

\[
A = \{(x_1, x_2, \ldots, x_n) : \inf\{k : x_1 + \ldots x_k = k - 1\} = n\} \\
B = \{(x_1, x_2, \ldots) : x_1 + \ldots x_k > k - 1 \forall k \in \mathbb{N}\}
\]

These are possible histories we may discover in our exploration algorithm, \( A \) corresponds to those histories which lead to extinction and \( B \) are those that never go extinct. We know that the probability of events in \( A \) given by:

\[
\mathbb{P}(H = (x_1, x_2, \ldots, x_n)) = \prod_{i=1}^{t} p_{x_i}
\]

(Recall here that \( p_x \) is the distribution function for \( X \), i.e. \( \mathbb{P}(X = x) = p_x \)). This is a probability distribution on \( A \) when we condition on the fact that the population eventually goes extinct. We will now condition on the fact that we want \( x_1 + \ldots + x_n = n - 1 \), and \( x_1 + \ldots + x_k > k - 1 \) for \( k < n \) to get the probability distribution we want.

Define a new birth distribution called the *conjugate* birth distribution that we denote with a \( \tau \):

\[
p'_x = \eta^{x-1} p_x
\]

This is a new probability distribution on the integers, for normalization is guaranteed by the fact that \( \eta = G_X(\eta) \):

\[
\sum_k p'_k = \sum_k p_k \eta^{k-1} = \eta^{-1} G_X(\eta) = \eta^{-1} \eta = 1
\]

Notice that the extinction probability for this process is 1, since:

\[
G_X(s) = \sum_{k=0}^{\infty} p_k \eta^{k-1} s^k = \eta \sum_{k=0}^{\infty} p_k (\eta s)^k = \eta^{-1} G_X(\eta s)
\]

So since \( \eta = G_X(\eta) \) is the smallest solution to that equation, the smallest solution to \( \psi = G_X(\psi) \) is:

\[
\psi = G_X(\psi) = \eta^{-1} G_X(\eta \psi) \\
\eta \psi = G_X(\eta \psi)
\]
Hence \( \eta \psi = \eta \), so \( \psi = 1 \). Hence \( \eta' \) which is the smallest root must be 1 too.

Even more importantly, the branching process defined by \( X' \) is exactly the branching process defined by \( X \) conditioned on the fact that extinction occurs:

\[
P(H = (x_1, \ldots, x_n) \mid \text{extinction}) = \frac{P(H = (x_1, \ldots, x_n) \cap \text{extinction})}{P(\text{extinction})} = \frac{P(H = (x_1, \ldots, x_n), n < \infty)}{\eta} = \eta^{-1} \prod_{i=1}^{n} p_{x_i}
\]

Now, since \((x_1, \ldots, x_n) \in A \) (since \( n < \infty \)) we know that \( \sum_{k=1}^{n} x_k = n - 1 \) so rewrite this as:

\[
P(H = (x_1, \ldots, x_n) \mid \text{extinction}) = \eta^{-1} \prod_{i=1}^{n} p_{x_i}
\]

Now we will return to the formula:

\[
P \{ T = n \} = \frac{1}{n} P \{ X_1 + X_2 + X_3 + \ldots + X_n = n - 1 \}
\]

\[
T = \inf \{ k : X_1 + X_2 + X_3 + \ldots + X_k = k - 1 \}
\]

This can be proven using the hitting time theorem below are rewriting \( Y_i = X_i - 1 \) and starting at \( x_0 = 1 \).

**Theorem 10.** Let \( \{ Y_i \} \) be i.i.d steps for a random walk starting at \( x_0 \geq 0 \), \( S_i = x_0 + \sum_{i=0}^{t} Y_i \). Suppose that:

\[
P( Y \geq -1 ) = 1
\]

Then the hitting time \( T_0 = \inf \{ n : S_n = 0 \} \) satisfies:

\[
P_{x_0} ( T_0 = n ) = \frac{x_0}{n} P_{x_0} ( S_n = 0 )
\]

(\( P_{x_0} \) here makes the dependence of the measure on the initial state \( S_0 = x_0 \) explicit, which will be needed in the proof). This is a crazy theorem!

**Proof.** (We prove it for every possible \( x_0 \geq 0 \) by induction on \( n \geq 1 \))

Base case \( (n = 1) \): If \( x_0 = 0 \) we have \( T_0 = 0 \), so both sides are 0 above. If \( x_0 = 1 \) then both sides are equal to \( P( Y_1 = 0 ) \). If \( x_0 \geq 2 \), since \( Y_{-1} \geq -1 \), both sides are zero again.

Induction step \( (n \geq 2 \) and assume holds for \( n - 1 \)): As in the base case, if \( x_0 = 0 \) both sides are zero, so assume \( x_0 \geq 1 \). Write:

\[
P_{x_0} ( T_0 = n ) = \sum_{s=1}^{\infty} P_{x_0} ( T_0 = n \mid Y_1 = s ) P ( Y_1 = s )
\]

\[
= \sum_{s=1}^{\infty} P_{x_0 + s} ( T_0 = n - 1 ) P ( Y_1 = s )
\]

\[
= \sum_{s=1}^{\infty} \frac{x_0 + s}{n - 1} P_{x_0 + s} ( S_{n-1} = 0 ) P ( Y_1 = s ) \text{ by induction hyp.}
\]
Now use \( P_{x_0 + s}(S_{n-1} = 0) = P_{x_0}(S_n = 0|Y_1 = s) \) to get:

\[
(n - 1)P_{x_0}(T_0 = n) = \sum_{s=-1}^{\infty} (x_0 + s)P_{x_0}(S_n = 0|Y_1 = s)P(Y_1 = s)
\]

\[
= \sum_{s=-1}^{\infty} (x_0 + s)P_{x_0}(S_n = 0 \cap Y_1 = s)
\]

\[
= x_0P(S_n = 0) + \sum_{s=-1}^{\infty} sE_{x_0}(1_{\{Y_1 = s\}1_{\{S_n = 0\}}})
\]

\[
= x_0P(S_n = 0) + E_{x_0}(Y_11_{\{S_n = 0\}})
\]

Now since \( Y_i \)'s are i.i.d we can write:

\[
E_{x_0}(Y_11_{\{S_n = 0\}}) = \frac{1}{n} \sum_{k=1}^{n} E_{x_0}(Y_k1_{\{S_n = 0\}})
\]

\[
= \frac{1}{n} E_{x_0}\left(\sum_{k=1}^{n} Y_k\right)1_{\{S_n = 0\}}
\]

\[
= \frac{1}{n} E_{x_0}(S_n - x_0)1_{\{S_n = 0\}}
\]

\[
= \frac{-x_0}{n} P_{x_0}(S_n = 0)
\]

So finally then:

\[
P_{x_0}(T_0 = n) = \frac{1}{n - 1} \left( x_0P_{x_0}(S_n = 0) + \frac{-x_0}{n} P_{x_0}(S_n = 0) \right)
\]

\[
= \frac{x_0}{n - 1} \left( 1 - \frac{1}{n} \right) P_{x_0}(S_n = 0)
\]

\[
= \frac{x_0}{n} P_{x_0}(S_n = 0)
\]

Which completes the induction step.

I skipped the next couple sections. \( \square \)

## 2 Configuration Model

The configuration model is a way to create a random graph with a prescribed degree sequence. Suppose we have a complex network of size \( n \) whose degrees are labeled with the set \([n] = \{1, 2, \ldots, n\}\) and the vertex \( i \) has degree \( d_i \). How can we create a random graph with this property?

We will create a random multigraph that has the above property. When we condition on the fact that the graph is simple, the resulting random graph will be uniform over all the possible graphs with the prescribed degree sequence.

### 2.1 Introduction

Fix \( n \) and let \( d = (d_i)_{i \in [n]} \). This will be the vector that stores the distribution of the degrees. We will suppose \( d_j \geq 0 \), because otherwise we could just as well delete node \( j \) from our graph. Let \( l_n = \sum_{i \in [n]} d_i \) be the sum of the degrees. Notice that \( l_n = 2|E| \) must be even because of the handshake lemma. Hence we can only hope to create a simple random graph when \( l_n \) is even.

**Definition 11.** Fix \( n, d = (d_i)_{i \in [n]} \), and \( l_n = \sum_{i \in [n]} d_i \). We will create a random multigraph with degree sequence \( d \) as follows. At every vertex \( j \), draw \( d_j \) half-edges leaving vertex \( j \) but not yet connected to anything. Number the half-edges \( 1.l_n \). Choose a pair of half-edges uniformly at random and connect them to make an edge between the vertices they are from. Repeat this until all the \( l_n \) half-edges are connected to something. (Notice they could leave some vertices connected to themselves, making a loop, or two vertices could be connected by multiple edges.) We call the resulting random graph \( CM_n(d) \).

An equivalent way to think about this is to label the half-edges as \( i \in l_n \) and then their partner is given by the bijection \( \sigma(i) : [l_n] \to [l_n] \) which has the property that \( \sigma(\sigma(i)) = i \) so that the configuration is stored in this bijection. Such a bijection is called a configuration.

**Example 12.** The number of possible configurations is \( (2l_n - 1)!! = (2l_n - 1)(2l_n - 3) \ldots 3 \cdot 1 \) since there are \( 2l_n - 1 \) choices available for \( \sigma(1) \), and then \( 2l_n - 3 \) choices for the next index which has not been chosen yet etc.

To go from the configuration to the configuration model graph, identify the set \( \{1, \ldots, d_1\} \) with vertex 1, \( \{d_1 + 1, \ldots, d_1 + d_2\} \) to vertex 2 etc.
Proposition 13. Let $G$ be a multi graph on $[n]$ and let $x_{ij} =$ #of vertices from $i$ to $j$. Let $d_i = x_{ii} + \sum_{j\in[n]}x_{ij}$ be the degree of vertex $i$. Then:

$$
P(C_{kn}(d) = G) = \frac{1}{(l_n-1)!} \prod_{i\in[n]}d_i! \prod_{1 \leq i < j \leq n}x_{ij}!.
$$

Proof. Since there are $(l_n-1)!$ configurations from which to choose from, (and they are uniformly selected), we have:

$$
P(C_{kn}(d) = G) = \frac{1}{(l_n-1)!} N(G)
$$

Where $N(G)$ is the number of configurations that give the graph. One can see this precisely $\prod_{i\in[n]}d_i! \prod_{1 \leq i < j \leq n}x_{ij}!$. There are $d_i!$ ways to permute the half-edges in $\{1, \ldots d_i\}$ which are identified with the vertex $1$, all of which will lead to the same outcome for the graph. Similarly for $d_i!$, so we see that $\prod_{i\in[n]}d_i!$ counts the relabeling possible of the half-edges without changing which vertex they correspond to. However, this counts some configurations more than once. Every loop leads to a double counting. (say the loop is from half edges $a, b$ which both correspond to vertex $1$. We count it once in $\{a, b\}$ and again in the rearrangement where $a$ and $b$ have switched spots $\{b, a\}$) Whenever there is more than one edge from $i$ to $j$ we will double count by a factor of $x_{ij}!$, for if we permute the $x_{ij}$ edges connecting 1 to 2 by the same permutation in $\{1, \ldots d_1\}$ and in $\{d_1 + 1, \ldots d_1 + d_2\}$ then we are back to the same configuration. So we divide by $\prod_{i\in[n]}2x_i! \prod_{1 \leq i < j \leq n}x_{ij}!$ to cancel these out.

Corollary 14. If $G$ is simple then $P(C_{kn}(d) = G) = const \delta$ so in particular, if $CM_{kn}(d)$ is simple, then it is a uniform selection from the set of all simple graphs with degree sequence $d_1, \ldots, d_n$.

Definition 15. Let $V$ be a vertex chosen uniformly at random from $[n]$ and let $D_n = \text{degree of vertex } V \text{ a random variable}$. Notice that the cumulative distribution function for $D_n$ is:

$$
F_n(x) = P(D_n \leq x) = \frac{1}{n} \sum_{j\in[n]} 1_{\{d_j \leq x\}}
$$

Notice that the only jumps for this CDF happen when $x$ an integer.

Definition 16. The following three regularity conditions will be useful:

I) Weak convergence of the distribution of vertex weights:

$$
D_n \Rightarrow D
$$

For some random variable $D$ distributed on the integers. (Equivalently, $F_n(x) \rightarrow F(x)$ for some $F$ and every $x \in \mathbb{R}$. This is equivalent here because $D_n$ is distributed on the integers, so for bounded continuous functions $f$ we have that $E(D_n f) = \sum_{x\in\mathbb{N}} f(x) P(D_n = x)$, so choosing bump functions with $f(x) = \begin{cases} 0 & x \neq k \\ 1 & x = k \end{cases}$ and the definition of weak convergence lets us see

$$
P(D_n = x) \rightarrow P(D = x) \text{ for every } x \text{, and similarly for } F_n.
$$

II) We will want $E(D_n) \rightarrow E(D)$ and $E(D_n^2) \rightarrow E(D^2)$. This is some sort of uniform integrability condition, see the notes on Billingsly.

Remark 17. Let $n_k = \sum 1_{\{d_i = k\}}$ be the number of vertices that have degree $k$. Up to relabeling of vertices, this uniquely determines $d$ (just write $n_1$ copies of “1”, $n_2$ copies of “2” etc.) Notice that $P(D_n = k) = \frac{n_k}{n}$, so the weak convergence condition above is that:

$$
\frac{n_k}{n} \rightarrow P(D = k)
$$

One canonical way to get this model is to choose $D$ up front and then generate the $n_k$'s by:

$$
n_k = \lfloor n P(D \leq k) \rfloor - \lfloor n P(D \leq (k-1)) \rfloor
$$

Notice here that $F_n(k) = \frac{1}{n} [nF(k)] \geq F(k)$ which is enough to give us uniform integrability for $D_n$ so we can see $E(D_n) \rightarrow E(D)$.

Remark 18. Another way to get this is to choose the sequence $d_1, \ldots, d_n$ to be an iid sequence of random variables $d_i \sim D$. Then $n_k = \sum 1_{\{d_i = k\}}$ is a sum of many iid Bernoulli random variables, and $\frac{n_k}{n} \rightarrow p_k$ is the Law of large numbers, so we have the weak convergence. Moreover, the random variable $D$ provides a bound that gives uniform integrability, so $E(D_n) \rightarrow E(D)$ too. Note that there is a problem here that $t_n$ might not be even, but we can always just add an extra half edge in this case that will not change things too much.