25.1 Galton-Watson Branching Process

The branching process, introduced by Galton and Watson in the 1870s, is a powerful tool to model population growth over discrete time. It has applications in the study of chemical chain reactions, scattering protons in nuclear fission, understanding genealogical structures, and many more.

Suppose at generation or time 0 there is only one individual who dies and is replaced at time 1 by a random number of offspring $X$. Next, suppose that these offspring dies and are themselves replaced at 2, each independently, by a random number of further offspring, having the same distribution as $X$, and so on. Let $Z_n$ denote the population size at time $n \geq 0$. The main problem is the probability of extinction, where no individuals exist after some finite number of generations.

Formally, let $(X_{n,i})_{n,i \geq 1}$ be i.i.d. non-negative integer-valued r.v.s with common distribution $X$. Define $Z_0 = 1$ and inductively, define

$$Z_n := \sum_{i=1}^{Z_{n-1}} X_{n,i}, \text{ for } n \geq 1.$$

We assume the following about the offspring distribution $X$,

i. PMF: $p_k = P(X = k), k \geq 0$,
ii. CDF: $F(k) = P(X \leq k), k \geq 0$, and
iii. MGF: $\varphi(s) = E(s^X) = \sum_{k=0}^{\infty} p_k s^k, s \in [0,1]$.

Clearly, the process $(Z_n)_{n \geq 0}$ is a Markov Chain with 0 being an absorbing state. We have $p_{ij} = P(Z_1 = j \mid Z_0 = i) = P(X_1 + X_2 + \cdots + X_i = j)$ for $i \geq 0, j \geq 0$.

Definition 25.1. The process $(Z_n)_{n \geq 0}$ is a Markov chain on $I = \{0,1,2,\ldots\}$ with 0 being an absorbing state and is known as the Galton-Watson Branching process with progeny distribution given by the distribution of $X$.

We can easily prove the following facts,

- if $p_1 = 1$, we have $Z_n = Z_0$ for all $n \geq 1$. 

25-1
• if $p_0 = 0, p_1 < 1$, then $Z_n \uparrow \infty$ as $n \to \infty$.
• the state 0 is absorbing.
• the limit $\rho := \lim_{n \to \infty} P(Z_n = 0 \mid Z_0 = 1)$ exists as $P_1(Z_n = 0) \leq P_1(Z_{n+1} = 0) \leq 1$. The number $\rho$ is called the extinction probability.

Define

$$\rho_n := P(Z_n = 0 \mid Z_0 = 1) = P_1(Z_n = 0), \quad n \geq 0.$$ 

Then we have the following,

$$\rho_{n+1} = P_1(Z_{n+1} = 0)$$ 
$$= P(\text{the population size at } n+1 \text{ generation is zero starting from one individual at generation 0})$$ 
$$= \sum_{k=0}^{\infty} p_k \cdot \rho_n^k = \varphi(\rho_n).$$

Hence, $\rho = \lim_{n \to \infty} \rho_n$ satisfies

$$\rho = \varphi(\rho),$$

where $\varphi$ is increasing and convex. Note that $\varphi(1) = 1, \varphi(0) = p_0 > 0$, and

$$\varphi'(1) = \sum_{k=0}^{\infty} k \cdot p_k = E(X) = \mu.$$ 

As illustrated in Figure 25.1,

• if $\mu \leq 1$, the only solution to $x = \varphi(x)$ is 1,
• if $\mu > 1$, there exists another solution to $x = \varphi(x)$.

![Figure 25.1: Plot of $x$ vs. $\varphi(x)$](image)

**Theorem 25.2.** For a Galton-Watson Branching Process (GWBP) with offspring distribution having mean $\mu$, we have the following:

i. If $\mu > 1$, the extinction probability is $< 1$ and is the minimum solution to $x = \varphi(x)$ (super-critical case),

ii. if $\mu < 1$, $\rho = 1$ (sub-critical case), and
Lecture 25: Branching Processes

iii. if $\mu = 1$ and $p_1 \neq 1, \rho = 1$ (critical case)

Check that, $E(Z_{n+1}) = \mu E(Z_n) = \cdots = \mu^{n+1}$. Thus if $\mu > 1$, with probability $(1 - \rho)$ the process will survive. Moreover, the growth is exponential under the condition $E(X \log^+ X) < \infty$.

25.1.1 Erdős–Rényi Random Graph $G(n, p)$

**Definition 25.3.** Define Erdős–Rényi random graph as a graph $G(n, p)$ with $n$ nodes, and $i \leftrightarrow j$ by an edge with probability $p$, independently for all $i < j$.

One can prove the following dichotomy when $p = \lambda/n$ for $\lambda \in (0, \infty)$,

- if $\lambda < 1$, all connected components are $\leq O(\log n)$, and
- if $\lambda > 1$, one giant component of size $\Theta(n)$ and every other components are $O(\log n)$.

Number of neighbors for a typical vertex $\sim \text{Binomial}(n - 1, p) \approx \text{Poisson}((n - 1)p)$ with mean $(n - 1)p \approx \lambda$ when $p = \lambda/n$. So the local growth can be approximated by a GWBP with Poisson($\lambda$) offspring distribution.

25.2 Markovian Branching Process

A *Continuous Time Branching Process* can be defined by replacing the lifetime of time 1 by i.i.d. continuous random variables. The main difference between discrete and continuous branching processes is that births and deaths occur at random times for continuous-time processes. Continuous time branching processes have the Markov property if and only if birth and death times are exponentially distributed. An example is shown in Figure 25.2.

We define a *(Markovian) continuous time binary branching process* by the following rules. Each individual gives birth to a new individual at rate $\lambda$ or dies at rate $\mu$. Individuals are independent of each other. More precisely, each individual in the population has two independent exponential random variables attached to it. One random variable $B$ has rate $\lambda$, the other one $D$ has rate $\mu$. If the rate $\lambda$ exponential random variable $B$ happens before the rate $\mu$ exponential $D$, then the individual is replaced by two individuals. If the rate $\mu$ exponential random variable $D$ happens before the rate $\lambda$ exponential $B$, then the individual dies with no offspring. Every new individual gets two independent exponential random variables attached to it (one with rate $\lambda$ and the other with rate $\mu$), and so on.

The number of individuals at time $t$ is denoted by $Z_t$. We start the process with $Z_0 = 1$ (see Figure 25.2). Then, $(Z_t)_{t \geq 0} \sim \text{Markov}(\delta_1, Q)$ on $I = \{0, 1, 2, \ldots\}$, with

$$Q = \begin{bmatrix}
0 & 0 & 0 & 0 & \cdots & \cdots \\
\mu & -(\lambda + \mu) & \lambda & 0 & \cdots & \cdots \\
0 & 2\mu & -2(\lambda + \mu) & 2\lambda & 0 & \cdots \\
\vdots & \vdots & \ddots & \ddots & \ddots & \ddots \\
\end{bmatrix}$$
where \( q_{ii} = i(\lambda + \mu), q_{i,i+1} = i\lambda, q_{i+1,i} = (i+1)\mu \) for \( i \geq 0 \).

The MGF of \( Z_t \) is

\[
F(t, s) = \mathbb{E}(s^{Z_t}) = \sum_{k=0}^{\infty} P_1(Z_t = k) \cdot s^k = \sum_{k=0}^{\infty} p_{1k}(t) \cdot s^k.
\]

The Backward Equation is

\[ P'(t) = QP(t) \text{ and } P(0) = I. \]

Thus,

\[
\frac{\partial F(t, s)}{\partial t} = \sum_{k=0}^{\infty} p_{1k}(t)s^k = \sum_{k=0}^{\infty} \left( \mu \cdot p_{0k}(t) - (\lambda + \mu) \cdot p_{1k}(t) + \lambda \cdot p_{2k}(t) \right) s^k
\]

\[
= \mu \cdot s^0 - (\lambda + \mu) \cdot F(t,s) + \lambda \cdot \sum_{k=0}^{\infty} p_{2k}(t) \cdot s^k,
\]

in which, \( \sum_{k=0}^{\infty} p_{2k}(t) \cdot s^k = \mathbb{E}_2(s^{Z_t}) = \mathbb{E}_{1,1}(s^{Z_t^{(1)}+Z_t^{(2)}}) = F(t,s)^2. \)

Thus, we have the following theorem,

**Theorem 25.4.** The function \( F(t, s), t \geq 0, s \in [0,1] \) satisfies

\[
\frac{\partial F(t, s)}{\partial t} = \mu - (\lambda + \mu)F + \lambda F^2, \text{ and } F(0,s) = s.
\]

**Extinction Probability.** If we define \( T_e \) as the extinction time, then we have the probability of extinction at time \( t \) as \( H(t) = P_1(T_e \leq t) = P_1(Z_t = 0) = F(t,0) \), with \( H(0) = 0 \) and

\[
H'(t) = \mu - (\lambda + \mu)H(t) + \lambda H(t)^2, \quad \forall \ t > 0.
\]

Thus, \( \rho = \lim_{t \to \infty} H(t) \) satisfies \( 0 = \mu - (\lambda + \mu)\rho + \lambda \rho^2 \Rightarrow \rho = 1, \frac{\mu}{\lambda} \). Based on the derivation above,

- if \( \mu < \lambda, \rho = \mu/\lambda < 1 \).
- if \( \mu \geq \lambda, \rho = 1, \text{ i.e., extinct with probability 1} \).

If \( \mu = 0 \), the above process is called a **Yule Process**.