In 1873 Francis Galton wrote an article asking for an understanding of how family names in England would go extinct. A year later, Henry William Watson gave a mathematical solution and together they wrote a math paper introducing the mathematical framework of the branching process.

Branching processes can be used to model:

- Philogenetic or family trees
- Atomic Chain Reactions
- BFS in a network
- Epidemics
- Rumor spreading
Galton-Watson Process

**Definition**

Let $\mu$ be a distribution on the non-negative integers. A Galton-Watson process with offspring distribution $\mu$ is a stochastic process with

$$Z_0 = 1$$

and

$$Z_n = X_{n,1} + X_{n,2} + \cdots + X_{n,Z_{n-1}}$$

where the sum is $Z_{n-1}$ independent rv's each with distribution $\mu$.

We think of $Z_n$ as the number of individuals in generation $n$. 
Exercise: Prove that $Z_n$ is a homogeneous Markov Chain.

Which states are recurrent? Which are transient?
A branching process ‘goes extinct’ if $Z_n = 0$ for some $n$.

We say a branching process ‘survives’ if it does not go extinct.
Say the offspring distribution has mean $\lambda$. What is $\mathbb{E}Z_n$?

Use conditioning.

$$
\mathbb{E}Z_n = \mathbb{E}[\mathbb{E}[Z_n|Z_{n-1}]]
= \mathbb{E}[\lambda Z_{n-1}]
= \lambda \mathbb{E}Z_{n-1}
$$

Then repeat the trick $n - 1$ more times.

$$
\mathbb{E}Z_n = \lambda^n
$$
Show that if $\lambda < 1$ then the branching process goes extinct with probability 1.
What if $\lambda = 1$ or $\lambda > 1$?
Let $y$ be the probability that a given branching process goes extinct. Let $p_k \Pr[X = k]$, the probability that one individual has $k$ offspring. Then:

$$y = \sum_{k=0}^{\infty} p_k y^k$$

Why?

This looks familiar:

$$y = G_x(y)$$

The Generating Function of the offspring distribution!
So we want to solve $y = G_X(y)$. There’s always one solution to this equation: $y = 1$. But is that the only solution?

[Plot of a Poisson distribution with different means]
What conditions on the offspring distribution imply there are multiple solutions?

Which solution is the correct extinction probability?
Draw a picture and use Taylor’s Theorem.

\[ G_X(1) = 1 \]
\[ G'_X(1) = \mathbb{E}X \]
\[ G_X(0) = p_0 \]

This shows that if \( \mathbb{E}X > 1 \) then there are multiple solutions. If \( \mathbb{E}X < 1 \), there is only the single solution \( y = 1 \). If \( \mathbb{E}X = 1 \), there is a single solution as long as \( p_0 > 0 \) (and if not, we know the branching process is just \( Z_n = 1 \) for all \( n \)).
So it remains to determine which of the solutions is correct for $\mathbb{E} X > 1$.

We start by computing the generating function of $Z_n$:

$$G_{Z_1}(s) = G_X(s)$$

$$G_{Z_2}(s) = \mathbb{E} s^{Z_2} = \mathbb{E}[\mathbb{E}[s^{Z_2}|Z_1]] = \mathbb{E}[G_X(s)^{Z_1}] = G_X(G_X(s))$$
And iterating,

\[ G_{Z_n}(s) = G_X(G_X(\cdots G_X(s))) \]

Let \( y_n = \Pr[Z_n = 0] \). Then

1. \( y_n \) is increasing in \( n \)
2. \( y_n \to y \) as \( n \to \infty \)
3. \( y_n = G_X(G_X(\cdots G_X(0))) \)
The probability of extinction of a branching process is the smallest non-negative root of the equation:

\[ y = G_x(y) \]

Proof: We have seen that the extinction probability satisfies the equation. Let \( \gamma \) be some non-negative root of the equation. We claim that \( y \leq \gamma \).
Extinction Probability

\[ y_1 = G_X(0) \leq G_X(\gamma) = \gamma \]

since \( G \) is a non-decreasing function.

\[ y_2 = G_X(y_1) \leq G_X(\gamma) = \gamma \]

and so on.. \( y_n \leq \gamma \) for all \( n \), and since \( y_n \to y \), we have

\[ y \leq \gamma \]

completing the proof.