

Branching Processes and Graphs

Emy Nguyen and Natasha Stapf

DRP Winter 2026

Abstract

Throughout the winter semester, we read and learned about Random Graph Theory. Within this paper, we will be introducing different mathematical concepts within Random Graph Theory and providing the reader with the appropriate background to be able to understand simple results and processes. We will be walking through different branching processes and proving several theorems we learned during these past few months, as well as providing examples to understand graphs and branching processes.

1 Introduction

The goal of this project was to gain an understanding of graphs, trees, forests, and graph theory with the intent of being able to understand how to apply it. In this report, we will provide the reader with enough background information to be able to gain a basic understanding of graph theory, probabilistic methods, and branching processes. Once that is acquired, we will then start explaining some more difficult processes and results that we learned about and proved during the semester. We will follow that up with an example of how graph theory can be applied to a real-world scenario and how it is a beneficial result.

2 Preliminaries

Before introducing branching processes, it is important to establish some foundational definitions of graph theory. This terminology is defined in the textbook, Graduate Texts in Mathematics, by Béla Bollobás, *Modern Graph Theory*.

2.1 Graphs

A graph G is defined as a pair (V, E) , where V is a set of vertices (or nodes) and E is a set of edges connecting pairs of vertices. Within this kind of graph, we can trace a path, which is a sequence of distinct vertices $v_0, v_1, \dots, v_n$ such that $(v_i, v_{i+1}) \in E$ for all $0 \leq i < n$. In the context of a branching process, a path represents a direct line of descent from an ancestor to its descendants.

If a path begins and ends at the same vertex, it forms a cycle. However, since an individual in a branching process cannot be their own ancestor, these processes are acyclic. By definition, a connected graph containing no cycles is known as a tree; we also refer to a tree as a connected acyclic graph. A disjoint union of trees is called a forest, where disjoint means that the graphs do not share any vertices, i.e., $V_1 \cap V_2 = \emptyset$, where V_i are the vertex sets.

A walk is defined as a sequence of vertices $v_0, v_1, \dots, v_k$ where each vertex is connected to the next by an edge. Unlike a path, a walk does not require the vertices to be distinct. It allows for the repetition of nodes and edges.

Since every individual in a branching process, except for the initial ancestor, has exactly one parent, the population organizes itself into this tree structure. This structure starts by the root, which is the first node/vertex of a graph (Z_0), the vertex representing the initial ancestor of the process. From the root, the process branches outward until it reaches the leaves, which are vertices with no children; these correspond to individuals who have zero offspring ($X = 0$). Furthermore, the tree allows us to visually group individuals by time or distance from the start; the set of vertices located at distance n from the root forms the n -th generation, denoted by Z_n .

2.2 Probability

A branching process is a stochastic model used to track the growth and evolution of populations. As noted by *Remco van der Hofstad*, these processes act as a bridge, connecting simple individual reproduction to the complex structures seen in large-scale random graphs [1].

The behavior of any branching process is given by a single discrete random variable, X , which represents the number of offspring produced by an individual. This variable is defined by three key components:

- **Probability Mass Function (PMF):** We denote $p_k = \mathbb{P}(X = k)$ as the probability that an individual has exactly k offspring. Since these are probabilities, they must sum to one: $\sum_{k=0}^{\infty} p_k = 1$.
- **Expectation ($\mathbb{E}[X]$):** This is the average number of offspring per individual:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k p_k$$

This is important to determine if a population grows or dies out.

- **Variance ($\text{Var}(X)$):**

$$\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

In the long-term, for any population, we use the Probability Generating Function (PGF). For a random variable X , the PGF is defined as:

$$G(s) = \mathbb{E}[s^X] = \sum_{k=0}^{\infty} p_k s^k$$

The PGF is important for calculating the extinction probability. We say a population becomes extinct if it eventually reaches a generation n where the number of individuals Z_n is zero. Hence, the PGF helps us determine if the lineage will disappear or survive indefinitely.

3 Core Concepts and Processes

3.1 Survival vs. Extinction

A branching process is the simplest model to show a population evolving in time. In the model we studied, we are assuming that each individual gives birth to a number of children with an independent and identical distribution (i.i.d.). In this model, here is some important notation to know before going into the theorems:

- $p_i = \mathbb{P}(\text{individual has } i \text{ children})$ – with the offspring distribution denoted $(p_i)_{i \geq 0}$.
- Z_n – is the number of individuals in the n^{th} generation, where $Z_0 = 1$; and it satisfies the recursion relation $Z_n = \sum_{i=1}^{Z_{n-1}} X_{n,i}$.
- $(X_{n,i})_{n,i \geq 1}$ is a doubly infinite array of i.i.d. random variables – where $X_{n,i}$ is the number of offspring of the i^{th} individual in the generation n .
- The law $(p_i)_{i \geq 0}$ of X is called the offspring distribution of the branching process.
- $\eta = \mathbb{P}(\exists n : Z_n = 0)$ – is the extinction probability.

Theorem 3.1 Survival vs. Extinction for Branching Processes

For a branching process with independent and identically distributed (i.i.d.) offspring X :

- I. $\eta = 1$ when $\mathbb{E}[X] \leq 1$
- II. $\eta < 1$ when $\mathbb{E}[X] > 1$
- III. $\eta = 1$ when $\mathbb{E}[X] = 1$ and $\mathbb{P}(X = 1) < 1$

The extinction probability η is the smallest solution in $[0, 1]$ of:

$$\eta = G_X(\eta) \quad (1)$$

With $s \mapsto G_X(s)$ being the probability generating function of the offspring distribution X , i.e.

$$G_X(s) = \mathbb{E}[s^X] \quad (2)$$

Before we are able to prove this theorem, we will go over a few key definitions we need to know:

- When $\mathbb{E}[X] \leq 1$, the population will die out with probability 1.
- When $\mathbb{E}[X] > 1$, there is a nonzero probability that the population will survive forever.
- The Probability Generating Function (PGF) of an integer-valued random variable X is the function $G_X(t) = \mathbb{E}[t^X]$, such that $G_X(0) = \mathbb{P}(X = 0)$.

Proof

First, we start with

$$\eta_n = \mathbb{P}(Z_n = 0). \quad (3)$$

Since $\{Z_n = 0\} \subseteq \{Z_{n+1} = 0\}$ by convention, which implies that $\eta_n \uparrow \eta$ (increases monotonically). Let

$$G_n(s) = \mathbb{E}[s^{Z_n}] \quad (4)$$

denote the PGF of Z_n , the number of individuals in the n^{th} generation.

Since $\mathbb{P}(X = 0) = G_X(0)$ for an integer-valued random variable X , we can conclude using the definition of a PGF that:

$$\eta_n = \mathbb{P}(Z_n = 0) = G_n(0) \quad (5)$$

Next, we condition on the first generation, using the fact that

$$\mathbb{P}(Z_1 = i) = \mathbb{P}(X = i) = p_i \quad (6)$$

For $n \geq 2$, by the law of total expectation:

$$\begin{aligned} G_n(s) &= \mathbb{E}[s^{Z_n}] \\ &= \sum_{i=0}^{\infty} p_i \mathbb{E}[s^{Z_n} \mid Z_1 = i] \\ &= \sum_{i=0}^{\infty} p_i G_{n-1}(s)^i \end{aligned}$$

Using the idea that each of the i individuals produce offspring in the n^{th} generation i.i.d., with the PGF equal to $s \mapsto G_{n-1}(s)$, and taking $G_X(s)$ for the generating function of the first generation, we derive:

$$G_n(s) = G_X(G_{n-1}(s)) \quad (7)$$

Now, to show that η_n satisfies the recurrence relation $\eta_n = G_X(\eta_{n-1})$, we plug in $s = 0$:

- Left side: $G_n(0) = \mathbb{P}(Z_n = 0) = \eta_n$.
- Right side: $G_X(G_{n-1}(0)) = G_X(\eta_{n-1})$.

Equating both sides, we find the recurrence relation $\eta_n = G_X(\eta_{n-1})$. Since $\eta_n \uparrow \eta$, it follows that:

$$\eta = G_X(\eta) \quad (8)$$

We have proved that η is a fixed point. All we have left to prove is that η is the smallest solution. Let $\psi \in [0, 1]$ satisfy $\psi = G_X(\psi)$. We claim that $\eta \leq \psi$ by induction:

Base case: $\eta_0 = 0$. Since $0 \leq \psi$, the base case holds.

Inductive step: Assume $\eta_{n-1} \leq \psi$. Since $s \mapsto G_X(s)$ is increasing on $[0, 1]$:

$$\eta_n = G_X(\eta_{n-1}) \leq G_X(\psi) = \psi \quad (9)$$

Therefore, by induction, $\eta_n \leq \psi$ for all n . Since $\eta_n \uparrow \eta$, we conclude that $\eta \leq \psi$ for all solutions ψ of $\psi = G_X(\psi)$. This tells us that η is the smallest solution.

3.2 The Random Walk

The Random Walk perspective is used when we are looking at random graphs. We use this perspective as it tends to be easier to construct a branching process by focusing on the number of children of each individual in the population rather than studying the number of descendants each individual has in a given generation, which is the more common way to construct a branching process.

To construct said representation of a branching process, we take $X'_1, X'_2, \dots$ to be i.i.d. random variables and we define $S'_0, S'_1, \dots$ to be the recursion

$$S'_0 = 1 \quad (10)$$

$$S'_i = S'_{i-1} + X'_i - 1 = X'_1 + \dots + X'_i - (i - 1) \quad (11)$$

Next, we take T' to be the smallest t such that $S'_t = 0$. If such a t does not exist, then we take $T' = +\infty$. This provides us with an equivalent definition of a branching process, but it presents the tree in a different way.

Breadth-first order example:

In breadth-first order, we take an active individual. And at time i we give it X'_i children, who are then marked as active, with the initial individual becoming inactive. This cycle continues as long as there continues to be active individuals in the population. The recursion we discussed above, S'_i describes the number of active individuals in the population after we have explored the first i individuals of the population. This process keeps going until we find that $S'_t = 0$. One caveat is that a branching process S'_t , only makes sense for $i \leq T'$ where $S'_i \geq 0$ only holds for $i \leq T'$, but the recursion itself can be defined for all $i \geq 0$.

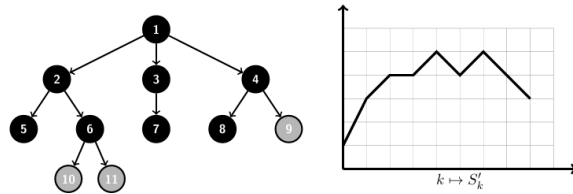


Figure 1: This diagram is an example of the process we talked about above. We explore the tree in breadth-first order, where we see that the black vertices have been explored, but the grey vertices have not been explored yet. Vertices 1-8 have been explored, $X_1 = 3, X_2 = 2, \dots$, and the next one to be explored is vertex 9. To the right of the graph, we have the plotted random-walk path we extracted from the exploration of the branching process tree.

Random Walks allow us to study the extinction probability for a tree when it has at least a given size. Theorem 3.8 is the main result to this process.

Theorem 3.8 Extinction Probability with Large Total Progeny

For a branching process with i.i.d. offspring X and mean $\mu = \mathbb{E}[X] > 1$,

$$\mathbb{P}(k \leq T < \infty) \leq \frac{e^{-Ik}}{1 - e^{-I}} \quad (12)$$

Where the exponential rate I is given by the formula

$$I = \sup_{t \leq 0} (t - \log \mathbb{E}[e^{tX}]) > 0 \quad (13)$$

This theorem states that when the total progeny, T , is large, then the branching process will survive with high probability.

3.3 Offspring Distribution

The Galton-Watson process is a mathematical model used to track how a population grows or dies out over time. We will focus on the two most common models for these offspring counts: the Binomial and the Poisson distributions.

3.3.1 Binomial

The Binomial distribution is used when there is a fixed limit on the number of offspring an individual can have.

Definition $X \sim \text{Binomial}(n, p)$ means we are looking at the number of successes in n independent trials. The probability that an individual has exactly k offspring is given by:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n \quad (14)$$

Here, $\binom{n}{k}$ represents the number of ways to choose k successful offspring out of n total possibilities.

Properties

- **Mean (μ):** $\mathbb{E}[X] = np$. This is the average number of offspring we expect. If $np \leq 1$, the population is likely to go extinct.
- **Variance (σ^2):** $\text{Var}(X) = np(1-p)$. This shows how much the number of offspring varies from the average.
- **PGF ($G(s)$):** The PGF for a Binomial variable is found by summing the series:

$$G(s) = \sum_{k=0}^n s^k \binom{n}{k} p^k (1-p)^{n-k} = (1-p+ps)^n \quad (15)$$

The PGF is useful because it allows us to handle the power of n trials easily when calculating multi-generational growth.

3.3.2 Poisson Offspring Distribution

The Poisson distribution does not have a fixed upper limit. It is used to model scenarios where the number of offspring could theoretically be very large, but they occur at a constant average rate, λ .

Definition When $X \sim \text{Poisson}(\lambda)$, the probability of having k offspring is:

$$\mathbb{P}(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots \quad (16)$$

This model assumes that while the chance of any specific "infection" or "birth" is small, the number of potential events is so large that the offspring count follows a smooth pattern.

Properties

- **Mean (μ):** $\mathbb{E}[X] = \lambda$. In this model, λ is the "critical value." If $\lambda > 1$, there is a chance the population survives forever.
- **Variance (σ^2):** $\text{Var}(X) = \lambda$.
- **PGF ($G(s)$):** The PGF for the Poisson distribution is particularly elegant:

$$G(s) = \sum_{k=0}^{\infty} s^k \frac{e^{-\lambda} \lambda^k}{k!} = e^{\lambda(s-1)} \quad (17)$$

The Poisson distribution is often used as an approximation for the Binomial distribution when n is very large and p is very small. A key feature is that its mean and variance are exactly the same.

4 Conclusion

By providing the basic definitions of graphs, paths, and cycles, we can now have a better understanding of population growth. To prove the critical survival versus extinction theorem, we have studied Galton-Watson branching processes, utilizing probability generating functions. This helped us demonstrate that the long-term fate of a population is fundamentally determined by its mean offspring number, $\mathbb{E}[X]$. By introducing the random walk, we showed how viewing a tree through its total progeny allows us to estimate the likelihood of survival on a large scale.

We have learned that a branching process serves as the simplest stochastic model for a population evolving over time, where individuals reproduce independently according to a fixed distribution. A common real-world application of this is found in epidemiology, where the offspring distribution represents the number of secondary infections caused by a single infected individual. Much like the survival conditions we proved, if the average number of secondary infections is less than or equal to one, the disease is destined for extinction. However, if it is greater than one, there is a non-zero probability of a large-scale outbreak.

Branching process models are very relevant in today's world because they allow us to make predictions about global outcomes such as the total size of a population or its risk of dying out. Although this project focused on basic branching processes, these concepts serve, nonetheless, as prerequisites for more advanced topics in graph theory. The results we have discussed can be seen in the study of Martingales, which provides a better way to analyze the growth rates of these processes. Furthermore, these tree-based approximations are the primary tool used to study the local behavior of complex networks in random graph models, such as the Erdős-Rényi graph.

The implications of these processes are everywhere. From the way diseases spread to how information moves through a network, branching processes give us a way to quantify uncertainty. They turn it into something we can predict and understand.

Acknowledgments

We want to thank our mentor, Catherine Fontaine, for leading us through the Directed Reading Program this winter and helping us learn how to read, present, and analyze math. Thank you to the organizers, Catherine Fontaine, Hazem Hassan, and Sasha Bell, for running the DRP this semester and for this amazing opportunity.

References

- [1] R. van der Hofstad, *Random Graphs and Complex Networks Volume 1*, Cambridge University Press, 2016
- [2] B. Bollobás, *Modern Graph Theory*, Graduate Texts in Mathematics, Springer