ABSTRACT

CONGDON, ELISABETH RAQUEL. The Asymptotic Behavior of Coalescence Time in Critical Branching Processes. (Under the Direction of Min Kang).

The random dynamics of a population evolving in time is a phenomenon worthy of mathematical study. From examining genetic drift and tracing genetic diseases to the recent increased interest in genetic ancestry reports, it is clear that there is a strong need for deepening and expanding our knowledge in this area. In this research, we study a specific type of stochastic process called a Markov process, for which the behavior of the current population does not depend on the full history of the process, but merely the previous step. Naturally, this type of stochastic process models the random creation and extinction of a species in time and in such case is called a branching process. In this research, we clarify and broaden previously found results relating to the so-called coalescence time for the branching process. The generation of the most recent common ancestor of the members of a generation belongs is called the total coalescence time for the generation in question, and it is yet another random variable. This definition has a natural extension to the pairwise coalescence time for any two randomly chosen individuals belonging to the generation in question.

We explore the asymptotic behavior of the coalescence time conditioned on specific observations imposed on the current generation. In particular, we consider the coalescence time under conditioning by multi-scaled rare events associated with the n-th generation. We conclude that certain events yield the same effect as conditioning on non-extinction, while others give rise to distribution functions that show previously unexplored behavior of the total coalescence time under these conditions. Additionally, we determine the effect of similar conditioning on the pairwise coalescence time for the process, computing the distribution function to which the conditioned process converges asymptotically.

The Asymptotic Behavior of Coalescence Time
in Critical Branching
Processes

by
Elisabeth Raquel Congdon

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Mathematics

Raleigh, North Carolina

2021

APPROVED BY:

Tien Khai Nguyen	Arvind Saibaba
Ralph Smith	Min Kang Chair of Advisory Committee

DEDICATION

To my father and mother for their constant love and support in my academic endeavors, and to my amazing husband who lifted my spirit when I was lacking courage in my abilities. Finally, to my daughter, Avery, and her sweet unborn sibling, so that they may know that all things are possible for them through the strength of God.

ACKNOWLEDGEMENTS

I would like to thank all those at North Carolina State University who have inspired my growth as a mathematician. To my dissertation advisor Dr. Kang for constantly encouraging me since the moment we met. You have guided me through the most challenging few years of my life, and you have demonstrated for me the passion required for a long and successful career in mathematics. Thank you also for being a caring and supportive presence during the times when I needed to put motherhood ahead of my mathematics work.

To my friends and professors at Northern Illinois University, thank you for starting me on this graduate school journey to my PhD years ago. I will always remember the mathematical foundation that I built during my Master's program and the influence it has had on where I am today.

To my friends and professors at Wheaton College, you inspire me to strive for excellence for the sake of Christ and His kingdom. Thank you for welcoming me with open arms as an instructor and thereby renewing my passion for undergraduate mathematics education.

Finally, a special and most heartfelt thank you to my husband, Stephen, and my parents, Dave and Melanie. I would not have come close to achieving my goals without your financial, physical, emotional, and spiritual support. Thank you for standing by me as I took the time needed to navigate through my graduate programs. Thank you for playing a huge role in raising Avery while I spent time studying, writing, and teaching. Your sacrifices and support mean the world to me.

TABLE OF CONTENTS

CHAPTER

1

INTRODUCTION

The pertinent definitions for this research are presented in Chapter 2. Sections 2.1 through 2.2 give rise to the definition of a branching process in Section 2.3 and subsequently the critical case considered in this research.

On the topic of branching processes in general, there has been a good deal of work done. The topic was first given a thorough treatment by T.E. Harris in [11] and has since been one of the major research topics in probability theory. The studies on the topic have been divided into three categories called the sub-critical, super-critical, and critical cases of the branching process. The essential distinction in these cases pertains to the concept of extinction. Let Z_n be a random variable representing the population size at time (generation) n of the process and Y_n a random variable representing the total coalescence time for the n-th generation. It can be shown that the expected number of offspring for a single individual influences the expected extinction behavior, in that with $Z_0 = 1$, when $\mathbb{E}(Z_1) \leq 1$, extinction will eventually occur in finite time with probability 1 and when $\mathbb{E}(Z_1) > 1$, extinction eventually occurs with probability $q \in (0,1)$. Details on this work follow in Section 2.3. These categories are further distinguished by more sophisticated differences between the cases where $\mathbb{E}(Z_1) < 1$ (sub-critical case) and $\mathbb{E}(Z_1) = 1$ (critical case). Other classifications of branching processes include time discrete or continuous, single-type or multi-type populations, and Markovian or non-Markovian dynamics. In this research, we focus on the critical case for the discrete branching process for single-type populations with Markovian dynamics.

Among the diverse literature of pertinent results to understanding such processes, there are a set of results in [1] that utilize well-known convergence results for the population size to explore the asymptotic behavior of the coalescence time in this case. The reader may refer to [2] to read about the asymptotic behavior of the coalescence time in the super-critical case. Several well-known results first proved by pioneers in this topic, Kolmogorov and Yaglom, are built upon. Kolmogorov in [15] found that while in the critical case $\mathbb{P}(Z_n > 0) \to 0$, the rate of this convergence is of the order n^{-1}. Further, Yaglom in [18] found that conditioning on non-extinction and then cutting down the process by a factor of n, we get a process that converges to a non-degenerate limit. These results are presented in Chapter 3 alongside a new result that explores the asymptotic behavior of the n-th population size in an unconditioned setting.

Chapter 4 shows the application of Zubkov of previously found results to the coalescence time. [3] indicates that in the critical case, the coalescence time is equally likely to have occurred in the recent past as it is to have occurred in the early generations of the process. This is not the case in the sub-critical or super-critical cases. Our research expounds on this by exploring the asymptotic behavior of the coalescence time conditioned on specific observations imposed on the current generation. In particular, we consider the coalescence time under conditioning by multi-scaled rare events associated with the n-th generation. We conclude that certain events yield the same effect as conditioning on non-extinction, while others give rise to distribution functions that show previously unexplored behavior of the total coalescence time. Section 4.2 considers a non-trivial adjustment to those mentioned above wherein we condition precise knowledge of the n-th population. For such consideration, we take advantage of the theory of quasi-stationary distributions, presented in Section 2.4.

Finally, we see in [1] that Athreya sought after an analogous result for the pairwise coalescence time for two randomly chosen individuals in the generation in question. His result in this case is presented in Chapter 5. Directly following his theory, we see that the pairwise coalescence time may be considered asymptotically under adjusted conditioning. The main result in this section precisely defines the distribution function to which the random variable representing pairwise coalescence converges asymptotically under the conditioning considered. To this end, the reader may see prerequisite results pertaining to point process theory in Section 2.5.1

CHAPTER

2

BACKGROUND MATERIAL

2.1 Random Variables

The outcomes of random experiments are expressed as the values of a function known as a random variable. The formal definition follows where $(\Omega, \mathscr{A}, \mathbb{P})$ is a probability triple with $\mathscr{A}$ a σ-algebra on Ω and $\mathbb{P}$ a probability measure on $\mathscr{A}$. All definitions in this section can be found in [12].

Definition 2.1.1. *A random variable is a function $X : \Omega \to \mathbb{R}$ such that it is measurable, i.e.,*

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathscr{A}$$

for all $x \in \mathbb{R}$.

In the case of a discrete random variable, X takes only values in some countable set $\{x_1, x_2, \ldots\}$. For the sake of this research, we will deal with random variables of the discrete type.

For a random variable, X, we may seek after its distribution function, which can be shown to always exist. First, we define the distribution function induced by a probability measure in the general sense.

Definition 2.1.2. *The distribution function induced by a probability measure* $\mathbb{P}$ *on* $(\mathbb{R}, \mathscr{B})$ *is the function* $F(x) : \mathbb{R} \to [0,1]$ *given by* $F(x) = \mathbb{P}((-\infty, x])$*. Here,* $\mathscr{B}$ *represents the Borel* σ*-algebra on* $\mathbb{R}$*.*

It can be shown that the distribution $F(x)$ uniquely characterizes the probability measure, as for any Borel set $B \subset \mathbb{R}$, $\mathbb{P}(B)$ may be obtained from the knowledge of F. As previously mentioned, it is useful to extend this notion to the random variable X.

Definition 2.1.3. *Let* X *be a random variable on* $(\Omega, \mathscr{A}, \mathbb{P})$ *taking values in* $(S, \mathscr{S})$ *(*$\mathscr{S}$ *is a* σ*-algebra on* S*). The distribution function of* X *(also called the law of* X*) is the function* $\mathbb{P}^X(B)$ *given by* $\mathbb{P}^X(B) = \mathbb{P}(X \in B)$ *for all* $B \in \mathscr{S}$*.*

It is natural to consider the distribution function of the law of X, denoted by $F_X(x)$ where

$$F_X(x) = \mathbb{P}^X((-\infty, x]) = \mathbb{P}(X \leq x).$$

This distribution function is particularly useful in studying random variables, because it, too, uniquely characterizes X and has certain useful properties. Note that with X acting as a discrete random variable, $F_X(x)$ is a jump function. We often call F_X the cumulative distribution function (CDF) of the random variable X.
It can be easily shown that the CDF satisfies the following:

- F_X is right continuous
- F_X is non-decreasing
- $\lim\limits_{x \to -\infty} F_X(x) = 0, \quad \lim\limits_{x \to \infty} F_X(x) = 1$

We now set out to define the process of integration (countable summation in the discrete case) against the law of a random variable. The expectation of a random variable is the weighted average of its outcomes. Namely,

Definition 2.1.4. *The expectation of the discrete random variable* X *taking values in* $(S, \mathscr{S})$ *is defined by*

$$\mathbb{E}(X) = \sum_{x \in S} x \mathbb{P}(X = x),$$

whenever the sum is absolutely convergent.

The expected value $\mathbb{E}(X)$ is often called the mean of X since it gives some intuition as to the expected average outcome of the random variable. If we wanted to consider the average distance from the mean for the outcome of a random experiment, we could compute the following, letting m denote the mean of X,

$$\mathbb{E}(|X-\mathbb{E}(X)|) = \mathbb{E}(|X-m|).$$

Instead of computing the above quantity, it is often easier to compute an adjusted expectation, defined in the following.

Definition 2.1.5. *The variance of the random variable X is defined to be*

$$\begin{aligned}
\sigma^2(X) &= \mathbb{E}((X-m)^2) \\
&= \mathbb{E}(X^2) - \mathbb{E}(2mX) + \mathbb{E}(m^2) \\
&= \mathbb{E}(X^2) - 2m^2 + m^2 \\
&= \mathbb{E}(X^2) - m^2, \\
&= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2,
\end{aligned}$$

where the linearity of expectation is used at leisure.

Often times, we need to consider the outcome of an experiment conditioned upon knowledge of a certain event occurring. For events A and B where $\mathbb{P}(B) > 0$, it is well known that the conditional probability of A given B is

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

In such cases, the conditional expectation of the random variable X can be defined. Here, it is given in a restrictive case, but nonetheless one that is sufficient for considering the applications to come.

Definition 2.1.6. *Let X and Y be random variables each taking countable many values $\{x_1, x_2, \dots\}$ and $\{y_1, y_2, \dots\}$ in $\mathbb{R}$ respectively. If $\mathbb{P}(X = x_j) > 0$, then*

$$\mathbb{E}(Y|X = x_j) = \sum_{k=1}^{\infty} y_k \mathbb{P}(Y = y_k | X = x_j).$$

As a final prerequisite for the computations to come, the following theorem allows for considering the expected value concerning function composition with a random variable as expectation against the law of the random variable. This is particularly useful when evaluating the expectation of a functional of a random variable with whose distribution function we are familiar.

Theorem 2.1.1 (Expectation Rule)**.** *Let X be a random variable on $(\Omega, \mathscr{A}, \mathbb{P})$ with values in $(S, \mathscr{S})$ and distribution $\mathbb{P}^X$. Let $f : (S, \mathscr{S}) \to (\mathbb{R}, \mathscr{B})$ be a measurable function. Then,*

(a) $f(X) \in \mathscr{L}^1(\Omega, \mathscr{A}, \mathbb{P})$ if and only if $f \in \mathscr{L}^1(S, \mathscr{S}, \mathbb{P}^X)$.

(b) If f is positive or if it satisfies (a), then

$$\int_\Omega f(X) d\mathbb{P}(x) = \int_\mathbb{R} f(x) d\mathbb{P}^X(x).$$

Equivalently, $\mathbb{E}_\mathbb{P}(f(X)) = \mathbb{E}_{\mathbb{P}^X}(f)$.

2.1.1 Convergence of a Sequence of Random Variables

Consider now a sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ defined on a common probability space. Since these are merely measurable functions, we seek after the definition of convergence in this context. There are various modes of convergence that are useful in probability theory including but not limited to pointwise convergence, almost sure convergence, $\mathscr{L}^p$ convergence, and convergence in probability. In a class of its own is the notion of weak convergence or convergence in distribution. In this case, it is not the random variables themselves that are shown to converge, but rather their associated sequence of probability distributions that is converging to a unique probability distribution. The theory and definitions regarding convergence of such random variables can be found in a variety of sources such as [12] or [4].

Definition 2.1.7. *Let $\{X_n\}_{n\in\mathbb{N}}$ and X be $\mathbb{R}^d$-valued random variables. X_n converges in distribution to X if the sequence of distribution functions $\mathbb{P}^{X_n}$ converges weakly to $\mathbb{P}^X$.*

As a result of this along with the result (2.1.1), it is immediate to see that this convergence in distribution described above occurs if and only if we have

$$\lim_{n\to\infty} \mathbb{E}(f(X_n)) = \mathbb{E}(f(X)),$$

for all test functions f on $\mathbb{R}^d$.

Probability theory takes advantage of numerous mathematical transforms that allow for computing desired results in a more advantageous space than that in which the problem was first posed. One such transform, defined in the following, is useful in many mathematical fields and lends itself well to considering the convergence of a sequence of random variables.

Definition 2.1.8. *Let μ be a probability measure on $\mathbb{R}^n$. Its Fourier transform is denoted $\widehat{\mu}$ and is a function on $\mathbb{R}^n$ given by*

$$\hat{\mu}(u) = \int e^{i\langle u,x\rangle} d\mu(x),$$

where $\langle u, x\rangle$ denotes the scalar product of $u, x \in \mathbb{R}^n$.

Each such Fourier transform of a probability measure can be shown to be a bounded continuous function satisfying $\widehat{\mu}(0) = 1$. With this general definition in mind, we define the characteristic function of a random variable by the Fourier transform against the law of the random variable.

Definition 2.1.9. *Let X be an $\mathbb{R}^n$-valued random variable. Its characteristic function φ_X defined on $\mathbb{R}^n$ is*

$$\varphi_X(u) = \int e^{i\langle u,x\rangle} d\mathbb{P}^X(x).$$

The following well-known theorem makes use of characteristic functions to re-frame convergence of probability measures in a way that is often more advantageous to compute.

Theorem 2.1.2 (Levy's Continuity Theorem)**.** *Let $\{\mu_n\}_{n\geq 1}$ be a sequence of probability measures on $\mathbb{R}^d$, and let $\{\widehat{\mu}\}_{n\geq 1}$ denote their Fourier transforms, or characteristic functions.*

(a) If μ_n converges weakly to a probability measure μ, then $\widehat{\mu}_n(u) \to \widehat{\mu}(u)$ for all $u \in \mathbb{R}^d$.

(b) If $\widehat{\mu}_n(u)$ converges to a function $f(u)$ for all $u \in \mathbb{R}^d$, and if in addition f is continuous at 0, then there exists a probability μ on $\mathbb{R}^d$ such that $f(u) = \widehat{\mu}(u)$, and μ_n converges weakly to μ.

Remark: The "weak" convergence of probability measures refers to convergence in distribution, whereas the convergence of characteristic functions is in the pointwise sense.

In its most general form, Levy's Continuity Theorem applies to sequences of probability measures. Let $\{X_n\}_{n\in\mathbb{N}}$, X be random variables on $(\Omega, \mathscr{A}, \mathbb{P})$ taking values in $(S, \mathscr{S})$ with associated distribution functions $\{\mathbb{P}_n^X\}_{n\in\mathbb{N}}$, $\mathbb{P}^X$. Applying Theorem 2.1.2 to this sequence of random variables, we see that the weak convergence of a sequence of random variables is essentially equivalent to the pointwise convergence of their associated characteristic functions.
In the study of convergence of sequences of random variables (To the end of applying this theory to point processes in Section 2.5.1), it is also helpful to consider the Laplace transform of a random variable.

Definition 2.1.10. *Let X be an $\mathbb{R}^n$-valued random variable. Its Laplace transform defined on $\mathbb{R}^n$ is*

$$\mathscr{L}(t) = \mathbb{E}(e^{-tX}),$$

$t \in \mathbb{R}^+$.

Recalling what weak convergence of random variables looks like from definition (2.1.7), we see immediately that weak convergence of a sequence of $\mathbb{R}^d$-valued random variables, $\{X_n\}_{n\in\mathbb{N}}$ to some random variable X, is equivalent to convergence of their Laplace transforms.

2.2 Markov Chains

A Markov Chain is a random process wherein the future of the chain is not dependent on the full history of the process, but rather on the current value. These processes are defined in terms of a sequence of random variables where the value of the random variable at a given time step gives the state of the process at that time. In the case of Markov chains, the index set is considered to be discrete (often representing time) and the state space, or set of possible outcomes, is assumed to be countable and discrete. The following definitions come from [9].

Definition 2.2.1. *The process $X = \{X_n \,:\, n \in N\}$ is a Markov chain if it satisfies the Markov property:*

$$\mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1}, \ldots, X_0 = x_0) = \mathbb{P}(X_n = x_n \mid X_{n-1} = x_{n-1})$$

for all $n \geq 1$ and for all $x_0, \ldots, x_n \in S$.

The conditional probabilities expressed in this definition describe the dynamics of the chain and are thus given a name and expounded upon to consider chain behavior as is carried out over a period of many "steps" of the process.

Definition 2.2.2. *The one-step transition probabilities for the process are given by*

$$\mathbb{P}(x, y) = \mathbb{P}(X_{n+1} = y \mid X_n = x)$$

for $x, y \in S$, and for any $m \geq 0$, the m-step transition probabilities are given by

$$\mathbb{P}^m(x, y) = \mathbb{P}(X_{n+m} = y \mid X_n = x)$$

for $x, y \in S$.

Homogeneity is a standard assumption in the study of such processes, which necessitates that the transition probabilities are stationary, i.e. unchanging in time. Considering p_{xy} as the xy-th entry in a potentially infinite dimensional matrix $\mathbf{P}$ (dimensioned $|S| \times |S|$), we obtain a matrix that contains the information on how this process moves from step to step in the process. This matrix $\mathbf{P}$ is often termed the transition matrix. By nature of its entries as values of a probability measure, it is straightforward to deduce that the following hold for $\mathbf{P} = (p_{xy})$:

- $p_{xy} \geq 0$ for all $x, y \in S$
- $\sum_y p_{xy} = 1$ for all $x \in S$

A consequence of the following theorem along with the homogeneity assumption is that

$$\mathbb{P}^m(x, y) = \mathbf{P}^m.$$

Theorem 2.2.1 (Chapman-Kolmogorov Equation)**.**

$$p_{ij}(m, m+n+r) = \sum_k p_{ik}(m, m+n) p_{kj}(m+n, m+n+r)$$

2.3 Branching Processes

A branching process is a special case of a Markov chain wherein the state of the process at n gives the population of a species. Naturally, the countable index set is taken to represent time, for which purpose we will take $n \in \mathbb{N}$. In this setting, the so-called "offspring" of generation n count towards the value of the chain at generation $n+1$. We make several natural and standard assumptions:

1. Each individual is assumed to produce offspring according to an identical distribution function. This distribution function is further assumed to give a nontrivial probability of producing more than one offspring, as well as not putting mass one on any particular state (or nonnegative integer).

2. Each individual is assumed to behave independently of one another.

3. The branching process begins at time 0 with an initial population size of 1 (An easily generalized assumption due to assumptions 1 and 2.)

Assumptions 1 and 2 above are considered in conjunction with one another by saying that the reproductions all occur in an i.i.d. fashion. The definition is made rigorous below and can be found in [3].

Definition 2.3.1. *A branching process* $\{Z_n,\ n \in \mathbb{N}\}$ *is a Markov chain for which* Z_n *represents population size of the* n*-th generation.*

- *Its offspring distribution is given by* $\{p_y\}_{y\geq 0} = \mathbb{P}(\xi_{n,i} = y)$ *where* $\{\xi_{n,i},\ i \geq 1,\ n \geq 0\}$ *is a family of independent random variables and for any* $n,\ i$, $\xi_{n,i}$ *denotes the number of offspring of individual* i *in generation* n.
- *Its transition function is given by* $\mathbb{P}(x, y) = p_y^{*x}$ *for* $y \geq 0$ *and* $x \geq 1$ *and* $\mathbb{P}(0, y) = \delta_{0y}$ *for* $y \geq 0$. *Here,* δ_{ij} *is the Kronecker delta and* $\{p_j^{*i}\}_{j\geq 0}$ *is the i-fold convolution of* $\{p_j\}_{j\geq 0}$.

It is clear to see that a recursive formulation of generation $n+1$ may be given by

$$Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} \xi_{n,i} & \text{if } Z_n > 0 \\ 0 & \text{if } Z_n = 0 \end{cases}.$$

A useful transform in the study of branching process is the probability generating function defined below for an arbitrary random variable.

Definition 2.3.2. *The probability generating function (PGF) for a random variable* X *is given by* $\mathbb{E}(t^X) = \sum_{j=0}^{\infty} t^j \mathbb{P}(X = j)$ *for all* $t \in \mathbb{R}$ *for which the sum converges.*

In the case of the branching process being considered, there are multiple PGFs we may look at. On its simplest level, the identical probability distribution itself, p_y, will be present in the PGF for the random variable $\xi_{n,i}$. In fact, the values $p_0, p_1, \ldots$ turn out to be coefficients of a power series in the representation that follows. In the sense that the distribution of each $\xi_{n,i}$ is identically given, the subscripts on the random variables are suppressed in the computation.

Definition 2.3.3. *The PGF for the offspring distribution is given by*

$$f(t) \doteq \mathbb{E}(t^{\xi}) = \sum_{j=0}^{\infty} \mathbb{P}(\xi = j) t^j = \sum_{j=0}^{\infty} p_j t^j = \sum_{j=0}^{\infty} \mathbb{P}(1, j) t^j,$$

for all $t \in \mathbb{R}$ *for which the sum converges.*

We then expect that since the offspring from generation n contribute to generation $n+1$, the PGF for the random variable Z_n, call it $f_{(n)}(t)$, may be expressed in terms of $f(t)$. Note that Z_n is a random sum of N of the random variables ξ where N is independent of

$\{\xi_i\}$. We compute:

$$\begin{aligned}
f_{(n)}(t) &= \mathbb{E}(t^{Z_n}) \\
&= \mathbb{E}(t^{\xi_{n-1,1}+\cdots+\xi_{n-1,Z_{n-1}}}) \\
&= \mathbb{E}(t^{\xi_{n-1,1}} \ldots t^{\xi_{n-1,Z_{n-1}}}) \\
&= \mathbb{P}(Z_{n-1}=0) + \sum_{k=1}^{\infty} \mathbb{E}(t^{\xi_{n-1,1}} \cdots t^{\xi_{n-1,Z_{n-1}}} | Z_{n-1}=k)\mathbb{P}(Z_{n-1}=k) \\
&= \sum_{k=0}^{\infty} (f(t))^k \mathbb{P}(Z_{n-1}=k) \\
&= f_{(n-1)}(f(t))
\end{aligned}$$

Thus we see that this generating function for the branching process is really just the iterated generating function for the offspring distribution;

$$f_{(n)}(t) \;=\; f_n(t) \;=\; \sum_{j=0}^{\infty} \mathbb{P}^n(1,j)t^j.$$

When the moments of these processes exist, they can be computed in terms of the PGF and its derivatives.

$$\mathbb{E}(Z_1) \;=\; \sum_{j=1}^{\infty} j p_j \;=\; \sum_{j=1}^{\infty} j\mathbb{P}(1,j) \;=\; f'(1) \doteq m,$$

the mean of the process. It then follows that $\mathbb{E}(Z_n) = \sum_{j=1}^{\infty} j\mathbb{P}^n(1,j) = m^n$. There is a trichotomy of long-term behavior in branching processes about the mean value; the cases $m<1$, $m=1$, or $m>1$ are termed the sub-critical, critical, and super-critical cases respectively.

We may also see that if $\sigma^2 = \mathrm{var}(Z_1)$, then

$$\mathrm{var}(Z_n) = \begin{cases} \frac{\sigma^2 m^{n-1}(m^n-1)}{(m-1)} & m \neq 1 \\ n\sigma^2 & m=1 \end{cases}.$$

There is a useful connection between the PGF for Z_n and the mean value m that has implications for the probability of the process going extinct. Aptly named, the extinction probability gives the probability of the chain reaching state 0 in finite time and thus can no longer produce any more offspring. In settings such as this, the state 0 is called an absorbing state because upon reaching this state, the chain has no chance of ever leaving it. We see that

$$\lim_{n\to\infty} f_n(0) = \lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \mathbb{P}(Z_i = 0, \text{ for some } i \geq 1) : \text{ the extinction probability.} \quad (2.1)$$

The limit above must exist because $f_n(t)$ is monotone decreasing in n and is bounded below by 0. In order for the chain to go extinct, one of two events must occur. Either the fist individual produces no offspring, or it produces k offspring that each themselves go extinct in finite time. Letting q denote the yet unknown extinction probability, this simple argument makes clear that

$$q = p_0 + \sum_{k=1}^{\infty} P(1,k) q^k = f(q)$$

and so q must be a fixed point of $f(t)$. By nature of $f(t)$ being strictly convex and increasing in $[0,1]$ and the knowledge that $f(0) = p_0$, $f(1) = 1$, q may be computed in the various cases.
Claim: If $\mathbf{m \leq 1}$, then $\mathbf{f(t) > t}$ for $\mathbf{t \in [0,1)}$.

Proof. $m = f'(1) \leq 1$. Assume to the contrary that $\exists \rho_0 \in [0,1)$ such that $f(\rho_0) = \rho_0$. Then by the Mean Value Theorem, $\exists c \in (\rho_0, 1)$ such that

$$f'(c) = \frac{f(1) - f(\rho_0)}{1 - \rho_0} = \frac{1 - \rho_0}{1 - \rho_0} = 1.$$

This is a contradiction of the fact that $f'(s) < f'(1) \leq 1$ for all $s \in (0,1]$ by the convexity of $f(t)$. □

Claim: If $\mathbf{m > 1}$, then $\mathbf{f(t) = t}$ has a unique root in $\mathbf{[0,1)}$.

Proof. We proceed with a proof by construction. We know that $f(0) - 0 = p_0 > 0$. Further, by Taylor's Theorem and since $f'(1) > 1$ in this case, $f(t^*) - t^* < 0$ for some $t^* \in (0,1)$ sufficiently close to 1. Therefore, we may apply the Intermediate Value Theorem to obtain $\rho_0 \in (0, t^*)$ such that $f(\rho_0) = \rho_0$. □

As a result of the claims above, in the sub-critical and critical cases, the only fixed point of $f(t)$ in $[0,1]$ occurs at the endpoint where $t = 1$ and hence the extinction probability must be 1. In the super-critical case, the following computation shows that the smaller such fixed point, $\rho \in (0,1)$, gives the probability of extinction. Letting q denote the extinction probability (and the limit of $f_n(0)$ as shown in (2.1)) and ρ denote any root of $f(t) = t$ on $[0,1]$,

$$0 \leq \rho \longrightarrow f_n(0) \leq f_n(\rho) = \rho \longrightarrow q \leq \rho.$$

It follows that $q = \rho$.
For more information about branching processes and related limit theorems, see [17] and [13].

2.3.1 The Critical Case

In particular, when $m = 1$, the branching process will almost surely go extinct in finite time by the previous remarks. For the purposes of this research, we mainly focus on the critical case, and thus we impose conditioning that attempts to "fix" this degenerate behavior. There is more than one way to adjust the critical branching process and avoid the fated extinction. The method that we employ in the work that follows deals with considering the same branching process Z_n, only conditioned on the event that the process is not extinct up until the current generation n. Hence, the branching process $\{Z_n\}$ under $\mathbb{P}_{\{Z_n>0\}}$ will be looked at more closely in the subsequent chapters.

2.4 The Quasi-Stationary Distribution

Analogous to the well-known notion of the stationary distribution for Markov chains, the quasi stationary distribution (QSD) is invariant under time evolution for the process conditioned on non-extinction. The theory that follows in this section is adapted from [7]. You may see [6] for more results pertaining to QSDs.
Let T_0 denote the time of extinction for the branching process. Because 0 is an absorbing state, $T_0 = \inf\{t > 0 \mid Z_t = 0\}$. We continue to let $S \subset \mathbb{R}^+$ denote the state space for the general Markov process. Prior to time T_0, the process takes values in $S^a \doteq S \setminus \{0\}$, the space of so-called allowed states. Recall that q denotes the probability of extinction in finite time, $q = P(T_0 < \infty)$. Further, let $\{\mathbb{P}_X,\ x \in S\}$ be the family of distributions with the initial condition $x \in S$.

Definition 2.4.1. *A probability measure ν on S^a is said to be a Quasi-Stationary Distribution (QSD) if for all measurable $B \subset S^a$, $n \in \mathbb{Z}^+$,*

$$\mathbb{P}_\nu(Z_n \in B \mid n < T_0) = \nu(B).$$

Equivalently,

$$\mathbb{P}_\nu(Z_n \in B) = \nu(B)\mathbb{P}_\nu(n < T_0),$$

where we use the fact that $\mathbb{P}_\nu(Z_n \in B,\ n < T_0) = \mathbb{P}_\nu(Z_n \in B)$ when $B \subset S^a$.
As usual, $\mathbb{P}_\nu = \int_{S^a} \mathbb{P}_x \, d\nu(x)$. This must be true for all t.

In the case of the discrete Markov chain that is the critical branching process, a QSD ν satisfies

$$\mathbb{P}_\nu(Z_n = x \mid n < T_0) = \nu(x),$$

for all $x \in S^a$, for all t. In the critical case of the branching process being considered, the population will eventually go extinct with probability one, i.e. $q = 1$, as seen in section (2.3). Once the population reaches 0, there is no more opportunity for reproduction and 0 is thus an absorbing state that is bound to be reached in finite time. This extinction time, however, may not occur for a very long time during which the population may exhibit some behaviors in which we are interested. This is precisely why we apply the study of QSD.
When one or more QSD can be shown to exist, they have certain useful properties, such as the following. Theorem 2.4.1 shows that when starting the process Z from a QSD ν, the killing time T_0 is geometrically distributed. It follows that the killing time T_0 has an exponential law. This implies that the rate of survival must be at most exponential for a QSD to exist (See Theorem 2.4.2.)

Theorem 2.4.1. *If ν is a QSD, then there exists $\alpha(\nu) \in [0,1]$ such that*

$$\mathbb{P}_\nu(n < T_0) = \alpha(\nu)^n$$

for all $n \in \mathbb{Z}^+$.

Proof. We obtain the following by the Markov property.

$$\begin{aligned}
\mathbb{P}_\nu(n < T_0) &= \mathbb{P}_\nu(Z_0 \neq 0, \ldots, Z_n \neq 0) \\
&= \mathbb{P}_\nu(Z_n \neq 0 \mid Z_0 \neq 0, \ldots, Z_{n-1} \neq 0)\mathbb{P}_\nu(n-1 < T_0) \\
&= \mathbb{P}_\nu(Z_n \neq 0 \mid Z_{n-1} \neq 0)\mathbb{P}_\nu(n-1 < T_0) \\
&= \mathbb{P}_\nu(Z_1 \neq 0 \mid Z_0 \neq 0)\mathbb{P}_\nu(n-1 < T_0) \\
&= \mathbb{P}_\nu(1 < T_0)\mathbb{P}_\nu(n-1 < T_0).
\end{aligned}$$

Thus,

$$\mathbb{P}_\nu(n-1 < T_0) = \frac{\mathbb{P}_\nu(n < T_0)}{\mathbb{P}_\nu(1 < T_0)},$$

finishing the proof. □

By induction, this theorem tells us that in particular, $\mathbb{P}_\nu(n < T_0) = \mathbb{P}_\nu(1 < T_0)^n$ and hence $\alpha(\nu) = \mathbb{P}_\nu(1 < T_0)$.
In Theorem 2.4.1 above, we saw that there exists $\alpha(\nu) \in [0,1]$ such that $\mathbb{P}_\nu(n < T_0) = \alpha^n$ for

all $n \in \mathbb{Z}^+$. Letting $\alpha(\nu) \in (0,1)$ to avoid trivialities, we may let $\theta(\nu) \doteq -\log \alpha(\nu) n$. Then,

$$\mathbb{P}_\nu(n < T_0) = e^{-\theta(\nu)n}$$

for all $n \in \mathbb{Z}^+$. We say that $\theta(\nu)$ gives the exponential rate of survival of ν.
Note that we may return to the definition of QSD and write an equivalent statement, namely that ν is a QSD if there exists $\theta(\nu) \in (0,\infty)$ such that for all measurable $B \subset S^a$, for all $n \in \mathbb{Z}^+$,

$$\mathbb{P}_\nu(Z_n \in B) \ = \ \nu(B) e^{-\theta(\nu)n}.$$

Furthermore, if ν is a QSD then for all $\theta < \theta(\nu)$ we have that $\mathbb{E}_\nu(e^{\theta n}) < \infty$.
To see this, recall that under $\mathbb{P}_\nu$, T_0 is geometrically distributed with parameter $\theta(\nu)$, that is, Then

$$\mathbb{E}_\nu(e^{\theta T_0}) \ = \ \int_{S^a} \mathbb{E}_x(e^{\theta T_0})\, d\,\nu(x) \ = \ \frac{1-e^{-\theta(\nu)}}{e^{-\theta} - e^{-\theta(\nu)}} < \infty.$$

The finiteness of this integral proves the assertion. This shows that in order for a QSD to exist, it is necessary that some exponential moments be finite. Let

$$\theta_x^* \doteq \sup\{\theta : \mathbb{E}_x(e^{\theta T_0}) < \infty\},$$

denote the exponential rate of survival of the process. Then for a QSD to exist, we must see that $\theta_x^* > 0$. This is formalized in the following:

Theorem 2.4.2. *The following equality holds:*

$$\theta_x^* \ = \ \liminf_{t\to\infty} -\frac{1}{t} \log \mathbb{P}_x(t < T_0).$$

Also a necessary condition for the existence of QSD is the existence of a positive exponential moment or equivalently, a positive exponential rate of survival: $\exists x \in S^a$ *such that* $\theta_x^* > 0$. *When this condition is met, we say that the process is exponentially killed.*

A final area of interest sufficient for our coverage of QSD is an interesting connection to linear algebraic topics. Let $\mathbf{P} = (p(x,y),\ x,y \in S)$ be the transition matrix for Z. Let $\mathbf{P}_a = (p(x,y) : \ x,y \in S^a)$ be the transition kernel restricted to the allowed states and denote by $\mathbf{P}_a^n$ the n-th power of $\mathbf{P}_a$ for $n \in \mathbb{Z}^+$. If ν is a probability distribution on S^a, denote by $\tilde{\nu} = (\nu(x) : \ x \in S^a)$ the associated row probability vector indexed by S^a. Then

$$\mathbb{P}_\nu(Z_n = y) = \tilde{\nu}\mathbf{P}_a^n(y).$$

In the following, let ν be a QSD. As a result of Theorem 2.4.1, we have for all $n \in \mathbb{Z}^+$,

$$\mathbb{P}_\nu(Z_n = y) = \tilde{\nu}\mathbf{P}_a^n(y)$$

$$\Longrightarrow$$

$$\mathbb{P}_\nu(n < T_0)\tilde{\nu} = \tilde{\nu}\mathbf{P}_a^n$$

$$\Longrightarrow$$

$$\alpha^n \tilde{\nu} = \tilde{\nu}\mathbf{P}_a^n$$

for some $\alpha \in (0,1)$. Equivalently,

$$\alpha \tilde{\nu} = \tilde{\nu}\mathbf{P}_a$$

and so $\tilde{\nu}$ is a left eigenvector of $\mathbf{P}_a$ with associated eigenvalue α.

2.4.1 The Yaglom Limit

Definition 2.4.2. *A probability measure μ on S^a is said to be a Yaglom Limit of the process if for all $y \in S^a$,*

$$\lim_{n\to\infty} \mathbb{P}_x(Z_n = y \mid n < T_0) = \mu(y)$$

$$\Longleftrightarrow$$

$$\lim_{n\to\infty} \frac{\mathbb{P}_x(Z_n = y)}{\mathbb{P}_x(n < T_0)} = \mu(y).$$

It is important to note that this limit quotient given above need not exist in general. Furthermore, this limit may sometimes coincide with the QSD (or one of the QSD) of the process, if there is one. The reader may refer to [16] and [5] to see several Yaglom-type limit results in the sub-critical and super-critical cases of the branching process respectively. We set out to find the Yaglom limit for the general branching process. A few helpful results lead up to the main Yaglom limit result in Theorem 2.4.6. Let

$$\tilde{b}_j \doteq \frac{\pi_j q^j}{\sum_{j=1}^{\infty} \pi_j q^j} \geq 0,$$

where q denotes the extinction probability as before. This quantity turns up in Theorem 2.4.6 below, but to see why, see the remarks made in the theorems to come, which are crucial to our research. These results regarding the Yaglom Limit, among others, can be found in [3].

Lemma 2.4.3.

$$f_n^{(j)}(t) \;=\; a_{n,j}(t) + f'[f_{n-1}(t)]f_{n-1}^{(j)}(t), \;\; n, j \geq 1,$$

where $a_{n,j}(t)$ is a power series with nonnegative coefficients.

Proof. The proof of this result is by induction on j.
Base case, $j = 1$:
Since $f_n(t) = f(f_{n-1}(t))$, by the chain rule for derivatives,

$$f_n'(t) \;=\; f'(f_{n-1}(t)) \cdot f_{n-1}'(t),$$

and so the result is true with $a_{n,1}(t) \equiv 0$.
Inductive step, assume true for some $k > 1$ and show true for $k+1$:
Assume that there exists a power series $a_{n,k}(t)$ with non-negative coefficients such that

$$f_n^{(k)}(t) \;=\; a_{n,k}(t) + f'(f_{n-1}(t))f_{n-1}^{(k)}(t).$$

Now we compute

$$\begin{aligned}
f_n^{(k+1)}(t) &= \left[f_n^{(k)}(t)\right]' \\
&= \left[a_{n,k}(t) + f'(f_{n-1}(t))f_{n-1}^{(k)}(t)\right]' \\
&= a_{n,k}'(t) + f'(f_{n-1}(t))f_{n-1}^{(k+1)}(t) + f_{n-1}^{(k)}(t)f''(f_{n-1}(t))f_{n-1}'(t),
\end{aligned}$$

and so the result is true with

$$a_{n,k+1}(t) \;=\; a_{n,k}'(t) + f_{n-1}^{(k)}(t)f''(f_{n-1}(t))f_{n-1}'(t).$$

□

Lemma 2.4.4. *Assume that $p_1 \neq 0$. Then*

$$\frac{\mathbb{P}^n(1,j)}{\mathbb{P}^n(1,1)} \uparrow \pi_j \leq \infty, \;\; j \geq 1.$$

Proof. To see that this ratio must converge as $n \to \infty$ to a well-defined (potentially infinite) limit, it is sufficient to prove that

$$\frac{\mathbb{P}^n(1,j)}{\mathbb{P}^n(1,1)}$$

is a non-decreasing sequence. Since $p_1 \neq 0$, $P^n(1,1) \neq 0$ for all n. In the following calculations, recall that $P^n(1,1) = f_n'(0)$ and with $f_n^{(j)}(t) = \sum_{j\geq 0}^{\infty} j!P^n(1,j)t^0$, we see that $P^n(1,j) = \frac{1}{j!}f_n^{(j)}(0)$.

We use Lemma 2.4.3 to compute

$$\begin{aligned}
\frac{\mathbb{P}^n(1,j)}{\mathbb{P}^n(1,1)} &= \frac{f_n^{(j)}(0)}{j!f_n'(0)} \\
&= \frac{a_{n,j}(0)+f'[f_{n-1}(0)]f_{n-1}^{(j)}(0)}{j!f'(f_{n-1}(0)f_{n-1}'(0)} \\
&\geq \frac{f'[f_{n-1}(0)]f_{n-1}^{(j)}(0)}{j!f'(f_{n-1}(0)f_{n-1}'(0)} \\
&= \frac{f_{n-1}^{(j)}(0)}{j!f_{n-1}'(0)} \\
&= \frac{\mathbb{P}^{n-1}(1,j)}{\mathbb{P}^{n-1}(1,1)}
\end{aligned}$$

□

Because of the probabilistic implications of results surrounding the sequence π_j, we will denote its PGF by $\mathscr{P}(t)$, that is,

$$\mathscr{P}(t) = \sum_{n=1}^{\infty} \pi_n t^n.$$

Theorem 2.4.5. *Assume that $p_1 > 0$. $\mathscr{P}(t)$ converges for $t \in [0,1)$. Furthermore, $\sum \pi_j$ converges if $m < 1$ and diverges if $m \geq 1$.*

This theorem has some very important implications for the ratio $\tilde{b}_j \doteq \frac{\pi_j q^j}{\sum_{j=1}^{\infty} \pi_j q^j}$. In the case $m = 1$, $q = 1$ and so with $\sum \pi_j$ diverging by Theorem 2.4.5,

$$\tilde{b}_j = 0 \text{ for } j \geq 1 \tag{2.2}$$

In the case $m \neq 1$, Theorem 2.4.5 concludes that

$$\sum \tilde{b}_j = 1 \text{ and } \tilde{b}_j > 0 \text{ for all } j \geq 1 \text{ such that } \mathbb{P}^n(1,j) > 0 \text{ for some } n \geq 1. \tag{2.3}$$

See notes within the proof of the following Theorem 2.4.6 to see the important implications of these observations on the Yaglom limit.

Theorem 2.4.6. *Assume $q > 0$ gives the extinction probability.*

(a) $\lim_{n\to\infty} \mathbb{P}(Z_n = j \mid Z_n > 0) = b_j$ *exists.*

(b) *If $m \neq 1$, the b_j is a probability function and its generating function $\mathscr{B}(t) = \sum_{j=0}^{\infty} b_j t^j$ is the unique solution to*

$$\mathscr{B}\Big(\frac{f(tq)}{q}\Big) = f'(q)\mathscr{B}(t) + (1 - f'(q)) \tag{2.4}$$

among generating functions vanishing at 0.

(c) *If $p_1 > 0$, then $b_j = \tilde{b}_j$ and $\mathscr{B}(s) = \frac{\mathscr{P}(qs)}{\mathscr{P}(q)}$.*

Remark: Part (b) gives a property that the generating function must hold in the noncritical cases. In the super-critical case, considering the Yaglom limit is not particularly interesting since $\lim_{n\to\infty} \mathbb{P}(Z_n > 0) = 1$. *Part (c) shows that in the critical case, the Yaglom limit is trivial, based on* (2.2) *above. That is,* $\lim_{n\to\infty} P(Z_n = j \mid Z_n > 0) = 0$ *when* $m = 1$.

Proof. In order to obtain that

$$\mathbb{P}(Z_n = k \mid Z_n > 0) \xrightarrow{n\to\infty} b_k,$$

for some sequence of constants $\{b_k\}_{k\in S}$, we would need to see this convergence for each $k \in S$, which is exhaustive. We turn to Levy's Continuity Theorem, Theorem 2.1.2, for an alternate approach. Consider $\mathbb{P}_{\{Z_n \mid Z_n > 0\}}$ and note that we seek to find its limit in the distributional sense.

$$\mathbb{P}_{\{Z_n \mid Z_n > 0\}} \xrightarrow{\mathscr{D}} B$$

for some B if and only if

$$\langle f, \mathbb{P}_{\{Z_n \mid Z_n > 0\}} \rangle \xrightarrow{n\to\infty} \langle f, B \rangle$$

for all $f \in \mathscr{C}_b$. Now taking $f = \chi_A$, this is equivalent to

$$\mathbb{P}(Z_n \in A \mid Z_n > 0) \xrightarrow{n\to\infty} B(A).$$

Finally, because our discrete state space is simply the disjoint union of singletons, we would obtain our goal, i.e. that

$$\mathbb{P}(Z_n = k \mid Z_n > 0) \xrightarrow{n\to\infty} b_k,$$

for some sequence of constants $\{b_k\}_{k\in S}$

In the following, we will make use of some already determined results regarding the generating function, $f(t)$, for the process $\{Z_n\}$ and its iterates. We set out to calculate the pointwise limit of $\mathbb{E}(t^{Z_n} \mid Z_n > 0)$, if it exists. For the following calculations, we fix t.

Case 1: m ≤ 1 and therefore q = 1.
Recall that $f_n(t) = \sum_{k=0}^{\infty} t^k \mathbb{P}(Z_n = k)$ and hence $f_n(0) = \mathbb{P}(Z_n = 0)$.

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}(t^{Z_n} \mid Z_n > 0) &= \lim_{n\to\infty} \sum_{j=0}^{\infty} t^j \mathbb{P}(Z_n = j \mid Z_n > 0) \\
&= \lim_{n\to\infty} \frac{1}{\mathbb{P}(Z_n > 0)} \sum_{j=0}^{\infty} t^j \mathbb{P}(Z_n = j, Z_n > 0) \\
&= \lim_{n\to\infty} \frac{1}{\mathbb{P}(Z_n > 0)} \sum_{j=1}^{\infty} t^j \mathbb{P}(Z_n = j) \\
&= \lim_{n\to\infty} \frac{\sum_{j=0}^{\infty} t^j \mathbb{P}(Z_n = j) - \mathbb{P}(Z_n = 0)}{1 - \mathbb{P}(Z_n = 0)} \\
&= \lim_{n\to\infty} \frac{f_n(t) - f_n(0)}{1 - f_n(0)} \\
&= 1 - \lim_{n\to\infty} \frac{1 - f_n(t)}{1 - f_n(0)},
\end{aligned}$$

It would be nice to know that this limit exists. How can we see that $G_n(t) \doteq \frac{1-f_n(t)}{1-f_n(0)}$ is increasing in n? We may define $G_n(t)$ recursively via

$$\begin{aligned}
G_n(t) &= \frac{1 - f_n(t)}{1 - f_n(0)} \\
&= \frac{1 - f(f_{n-1}(t))}{1 - f_{n-1}(t)} \cdot \frac{1 - f_{n-1}(0)}{1 - f(f_{n-1}(0))} \cdot \frac{1 - f_{n-1}(t)}{1 - f_{n-1}(0)} \\
&= \frac{1 - f(f_{n-1}(t))}{1 - f_{n-1}(t)} \cdot \frac{1 - f_{n-1}(0)}{1 - f(f_{n-1}(0))} \cdot G_{n-1}(t).
\end{aligned}$$

Therefore, taking $H(t) \doteq \frac{1-f(t)}{1-t}$, this relationship can be expressed as

$$G_n(t) = \frac{H(f_{n-1}(t))}{H(f_{n-1}(0))} \cdot G_{n-1}(t).$$

$H(t)$ is non-decreasing in t, as is $f_n(t)$ and therefore $G_n(t)$ is non-decreasing in n. Denote by $G(t)$ the limit as $n \to \infty$ of $G_n(t)$. This proves that $\lim_{n\to\infty} \mathscr{B}_n(t) = 1 - G(t) = \mathscr{B}(t)$ exists. This proves (i). We now may note that

$$G_n(f(t)) = G_{n+1}(t) H(f_n(0)). \tag{2.5}$$

As $n \to \infty$, $f_n(0) \to q = 1$ and so we consider $\lim_{x\to 1} H(x) = \lim_{x\to 1} \frac{1-f(x)}{1-x} = f'(1) = m$. So, taking limits as $n \to \infty$ on both sides of (2.5) yields

$$G(f(t)) = mG(t), \tag{2.6}$$

or

$$\mathscr{B}(f(t)) = m\mathscr{B}(t) + (1-m). \tag{2.7}$$

We now restrict the current case to $m < 1$. Letting $\gamma = f'(q)$, with $q = 1$, $\gamma = m$. Thus,

$$\mathscr{B}(\frac{f(tq)}{q}) = \gamma\mathscr{B}(t) + (1-\gamma). \tag{2.8}$$

To see that $\{b_j\}$ is a probability function, it remains to see that $\lim_{t\to 1^-} \mathscr{B} = 1$. Well, taking $t \to 1^-$ in (2.6), $\lim_{t\to 1^-} G(f(t)) = \lim_{t\to 1^-} mG(t)$ and so with $m < 1$, $\lim_{t\to 1^-} G(t) = 0$ and thus $\lim_{t\to 1^-} \mathscr{B} = 1$. The last detail needed in part (ii) is that of uniqueness. Suppose that $\mathscr{R}$ is another solution to (2.4). Then both $\mathscr{B}$ and $\mathscr{R}$ satisfy (2.4) as well as an iterative analogous equation,

$$\mathscr{B}[f_n(t)] = \gamma^n \mathscr{B}(t) + (\gamma^{n-1} + \cdots + \gamma + 1)(1-\gamma). \tag{2.9}$$

Differentiating (2.9) yields

$$\mathscr{B}'[f_n(t)]f_n'(t) = \gamma^n \mathscr{B}'(t). \tag{2.10}$$

Now for any $t \in [0, q)$, there is a k such that $f_k(0) \le t \le f_{k+1}(0)$ and hence by (2.10),

$$\frac{\mathscr{R}'(s)}{\mathscr{B}'(t)} = \frac{\mathscr{R}'(f_n(t))}{\mathscr{B}'(f_n(t))} \le \frac{\mathscr{R}'(f_{n+k+1}(0))}{\mathscr{B}'(f_{n+k}(0))} = \frac{\mathscr{R}'(f_{n+k+1}(0))}{\mathscr{B}'(f_{n+k+1}(0))} \cdot \frac{\mathscr{B}'(f_{n+k+1}(0))}{\mathscr{B}'(f_{n+k}(0))},$$

but also

$$\frac{\mathscr{R}'[f_n(0)]}{\mathscr{B}'[f_n(0)]} = \frac{\mathscr{R}'(0)}{\mathscr{B}'(0)} = \text{constant}, \quad n \ge 1.$$

Hence,

$$\begin{aligned}\frac{\mathscr{R}'(t)}{\mathscr{B}'(t)} &\le \frac{\mathscr{R}'(0)}{\mathscr{B}'(0)} \cdot \frac{\mathscr{B}'(f_{n+k+1}(0))}{\mathscr{B}'(f_{n+k}(0))} \\ &= \frac{\mathscr{R}'(0)}{\mathscr{B}'(0)} \cdot \frac{f'_{n+k}(0)}{f'_{n+k+1}(0)}\gamma \\ &= \frac{\mathscr{R}'(0)}{\mathscr{B}'(0)} \cdot \frac{\gamma}{f'(f_{n+k}(0))}.\end{aligned}$$

Letting $n \to \infty$, we see that $\frac{\mathscr{R}'(t)}{\mathscr{B}'(t)} \le \frac{\mathscr{R}'(0)}{\mathscr{B}'(0)}$. The converse inequality holds similarly and so since $\mathscr{R}(0) = \mathscr{B}(0)$, $\mathscr{R} \equiv \mathscr{B}$. This completes the proof of (ii) in this case. We turn our attention towards part (iii). Recall some important quantities:

$$\mathscr{P}(t) = \sum_{n=1}^{\infty} \pi_n t^n \text{ and } \tilde{b}_j = \frac{\pi_j q^j}{\sum_{j=1}^{\infty} \pi_j q^j}.$$

Recall that $\lim_{n\to\infty} \mathbb{P}(Z_n = j|Z_n > 0) = b_j$. Wishing to express $\mathbb{P}(Z_n = j|Z_n > 0)$ in terms of π_j, we see that

$$\mathbb{P}(Z_n = j|Z_n > 0) = \frac{\mathbb{P}(Z_n = j, Z_n > 0)}{\mathbb{P}(Z_n > 0)} = \frac{\mathbb{P}^n(1,j)q^j}{\sum_{k=1}^{\infty} \mathbb{P}^n(1,k)q^k}.$$

By Lemma 2.4.4,

$$\frac{\mathbb{P}^n(1,j)q^j}{\sum_{k=1}^{\infty} \mathbb{P}^n(1,k)q^k} = \frac{\left[\frac{\mathbb{P}^n(1,j)}{\mathbb{P}^n(1,1)}\right]q^j}{\sum_k \left[\frac{\mathbb{P}^n(1,k)}{\mathbb{P}^n(1,1)}\right]q^k} \to \frac{\pi_j q^j}{\sum_k \pi_k q^k} = \tilde{b}_j.$$

We have proven that $b_j = \tilde{b}_j$. (2.3) and (2.2) tell precisely what the implications of this equivalence are in the three cases of mean offspring behavior. Finally,

$$\frac{\mathscr{P}(qt)}{\mathscr{P}(q)} = \frac{\sum_{n=1}^{\infty} \pi_n q^n t^n}{\sum_{n=1}^{\infty} \pi_n q^n} = \sum_{j=1}^{\infty} b_j t^j = \mathscr{B}(t).$$

Case 2: m $>$ 1 and therefore q $\in (\mathbf{0}, \mathbf{1})$.
Similarly to case 1,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}(t^{Z_n} \mid Z_n > 0) &= \lim_{n\to\infty} \frac{1}{\mathbb{P}(Z_n = 0)} \sum_{j=1}^{\infty} t^j \mathbb{P}(Z_n = j) \\
&= \frac{\sum_{j=1}^{\infty} t^j q^j \mathbb{P}(Z_n = j)}{\sum_{j=1}^{\infty} q^j \mathbb{P}(Z_n = j)} \\
&= \frac{f_n(tq) - f_n(0)}{q - f_n(0)}.
\end{aligned}$$

Let $g_n(t)$ be the n-fold iterate of $g(t) \doteq \frac{f(qt)}{q}$. Then

$$\frac{f_n(tq) - f_n(0)}{q - f_n(0)} = \frac{g_n(t) - g_n(0)}{1 - g_n(0)},$$

where $g_n(t)$ has mean $m_g = g'(1) = f'(q) = \gamma < 1$. We may apply the reasoning from case

1 to obtain the result (i). (2.4) also follows directly from substituting in g for f in (2.7), completing the proof of part (ii). Finally, part (iii) is the same as in case 1. $\square$

2.5 Point Processes

Put simply, a point process is a random cloud of points. Let $E \subseteq \mathbb{R}^n,\ n \in \mathbb{N}$ be a locally compact topological space with a countable basis and $\mathscr{E} = \mathscr{B}(E)$, its Borel σ-algebra. The prerequisite material presented below on the topic of point processes comes from [8]. See also [14] for additional results pertaining to related processes.

Definition 2.5.1. *A point measure m on E is a measure such that*

$$m = \sum_{i=1}^{\infty} \delta_{x_i},$$

where δ_{x_i} is the dirac measure at $x_i \in E$.

Definition 2.5.2. *A measure m is said to be a Radon measure if $m(K) < \infty$ for all $K \in \mathscr{E}$ compact, and m is both inner regular and outer regular.*

Further, let $M_p(E)$ be the space of all Radon point measures on E.

Definition 2.5.3. *A point process on E is a measurable map from a probability space Ω to $M_p(E)$ equipped with its σ-algebra $\mathscr{M}_p(E)$. $\mathscr{M}_p(E)$ is the smallest σ-algebra containing $\{m \in M_p(E) : m(F) \in B\}$ for all $F \in \mathscr{E}$ and $B \subset [0,\infty]$ Borel.*

For a point process, N, it is common to define the notion of the intensity of the process via the following.

Definition 2.5.4. *The function $\mu : \mathscr{E} \to \mathbb{R}^+ \cup \{\infty\}$ given by*

$$\mu(B) = \mathbb{E}(N(B)),$$

for all $B \in \mathscr{E}$ is a measure called the intensity measure for the point process N.

One common extension of point process theory is a study into the so-called Poisson point process, wherein the intensity measure of N has a Poissonian distribution with parameter $\mu(B)$.

2.5.1 Convergence of Point Processes

Let $\mathbb{N}_0 \doteq \mathbb{N} \cup \{0\}$ and let $\{N_n\}_{n\in\mathbb{N}_0}$ be a sequence of point processes.

Definition 2.5.5. *We say that N_n converges weakly to N_0, if for all bounded, continuous, real-valued functions f,*

$$\lim_{n\to\infty} \int_{M_p(E)} f \; d(N_n)_\star \mathbb{P} = \int_{M_p(E)} f \; d(N_0)_\star \mathbb{P},$$

where $(N_n)_\star \mathbb{P}$ is the push-forward measure of $\mathbb{P}$ by N_n for each n. This measure is clearly defined on $\mathscr{M}_p(E)$.

Section (2.1.1) affirms that weak convergence of a sequence of random variables is equivalent to converge of their Laplace transforms. This concept lends itself well to the study of point processes, as taking the Laplace transforms "smooths out" the process and creates a more manageable object of which to consider the convergence.

Definition 2.5.6. *For a point process N, the Laplace functional of N is the operator*

$$\mathscr{L}_N(f) = \mathbb{E}(e^{-N(f)}),$$

where $N(f) = \langle f, N\rangle = \sum_{x\in supp(N)} f(x)$ with f a non-negative function with bounded support.

CHAPTER

3

ASYMPTOTIC BEHAVIOR OF THE POPULATION SIZE IN THE CRITICAL CASE

In this chapter, we begin our investigation into the asymptotic behavior of the coalescence time by first making several observations about the random variable Z_n and how its behavior is dictated by the mean of the process. To that end, we recall from Definition 2.3.1 that the random variable Z_n represents the population size of the n-th generation. Section 2.3 goes on to define several quantities that will appear in this chapter. In particular, $m \doteq \mathbb{E}(Z_1)$ is introduced and explained to give the mean of the process $\{Z_n,\ n \in \mathbb{N}\}$. The discussion and results to follow will always consider the critical case of the branching process, wherein $m = 1$.

Theorems 3.1.1 and 3.2.1 are previously known results pertaining to the asymptotic behavior of the process. One considers the asymptotic behavior of $\{Z_n\}$ and the other considers the process $\{\frac{Z_n}{n}\}$ under $\mathbb{P}_{\{Z_n>0\}}$. Theorem 3.2.2 is an original result which considers the asymptotic behavior of $\{\frac{Z_n}{n}\}$, but not under any conditioning.

3.1 Without Conditioning

It is clear that in the critical case, $\mathbb{P}(Z_n > 0)$ will converge to 0 as $n \to \infty$, but one may be interested in the speed of such a convergence. The following result from [3] explores this notion.

Theorem 3.1.1. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Let* $Z_0 = k < \infty$. *Then,*

$$\lim_{n\to\infty} n\mathbb{P}(Z_n > 0) = \frac{2}{\sigma^2}.$$

Remark: We know that with $m = 1$ *the process goes extinct eventually with probability* 1*, but this theorem tells us just how fast. We see that* $\mathbb{P}(Z_n > 0) \sim \frac{2}{\sigma^2 n}$.

Proof. Recall that $f_n(t) = \mathbb{E}(t^{Z_n})$, so that $f_n(0) = \mathbb{P}(Z_n = 0)$ and $\mathbb{P}(Z_n > 0) = 1 - f_n(0)$. Suppose that

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{1-f_n(t)} - \frac{1}{1-t}\right) = \frac{\sigma^2}{2}, \tag{3.1}$$

uniformly for $t \in [0,1)$. Then in particular, when $t = 0$ we have

$$\frac{\sigma^2}{2} = \lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{1-f_n(0)} - 1\right) = \lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{\mathbb{P}(Z_n > 0)} - 1\right).$$

Thus we may turn our attention to proving (3.1) to obtain the desired result.
By L'Hopital's Rule,

$$\begin{aligned} \lim_{t\to 1^-} \frac{f(t)-t}{(1-t)^2} &= \lim_{t\to 1^-} \frac{f'(t)-1}{2t-2} \\ &= \lim_{t\to 1^-} \frac{f''(t)}{2} \end{aligned}$$

Now the generating function for the offspring distribution, f, is assumed to have finite variance, denoted by σ^2. Because of this assumption and its correspondence to the regularity of f, the limit above is well-defined. We may then let a be a finite real number such that $a \doteq \frac{\sigma^2}{2}$, and so, $\lim_{t\to 1^-} \frac{f''(t)}{2} = a$.
Let $\epsilon(t) \doteq a - \frac{f(t)-t}{(1-t)^2}$. Note that $a = \lim_{t\to 1^-} \dfrac{f(t)-t}{(1-t)^2}$ and further that $f(t)$ is strictly increasing. Then the rational function $\frac{f(t)-t}{(1-t)^2}$ is also strictly increasing to its limit a from the left-hand side. (Note that $f(t) - t > 0$ for $t \in [0,1)$ in the critical case as was proven in Section 2.3). Then, $a \geq \frac{f(t)-t}{(1-t)^2}$, and so $\epsilon(t) \geq 0$. Further, $\lim\limits_{t\to 1^-} \epsilon(t) = 0$.
Let $\delta(t) \doteq a - \left(\frac{1}{1-f(t)} - \frac{1}{1-t}\right) = a - \left(\frac{f(t)-t}{(1-f(t))(1-t)}\right)$. Since $t \leq f(t)$ and $\delta(t) \leq \epsilon(t)$, we sum

over i and replace t by $f_i(t)$ to obtain

$$\sum_{i=0}^{n-1} \delta(f_i(t)) \leq \sum_{i=0}^{n-1} \epsilon(f_i(t)).$$

We may note that $f_i(t)$ increases to 1 and so $\epsilon(f_i(t))$ decreases to 0 on $[0,1)$. Then

$$\sum_{i=0}^{n-1} \epsilon(f_i(t)) \leq \sum_{i=0}^{n-1} \epsilon(f_i(0)) \sim o(n).$$

We now need to look closer at $\sum_{i=0}^{n-1} \delta(f_i(t))$. We have

$$\begin{aligned}
\delta(s) &= a - \left(\frac{1}{1-f(s)} - \frac{1}{1-s}\right) \\
&= \left(a \cdot \frac{1-f(s)}{1-s} - \frac{f(s)-s}{(1-f(s))(1-s)} \cdot \frac{1-f(s)}{1-s}\right)\left(\frac{1-s}{1-f(s)}\right) \\
&= \left(a \cdot \frac{1-f(s)}{1-s} - \frac{f(s)-s}{(1-s)^2}\right)\left(\frac{1-s}{1-f(s)}\right) \\
&\geq a \cdot \frac{s-f(s)}{1-f(s)},
\end{aligned}$$

since $\frac{f(s)-s}{(1-s)^2} \leq a$.

$$\begin{aligned}
&= -a \cdot \frac{f(s)-s}{1-s} \cdot \frac{1-s}{1-f(s)} \\
&= -a(1-s)(a-\epsilon(s)) \cdot \frac{1-s}{1-f(s)},
\end{aligned}$$

since $\frac{f(s)-s}{1-s} = (1-s)\frac{f(s)-s}{(1-s)^2} = (1-s)(a-\epsilon(s))$.

$$\begin{aligned}
&\geq -a^2(1-s) \cdot \frac{1-s}{1-f(s)} \\
&\geq -a^2(1-s)\frac{1}{1-f(0)},
\end{aligned}$$

since $\frac{1-s}{1-f(s)}$ is decreasing and positive on $[0,1]$.

So,
$$\sum_{i=0}^{n-1}\delta(f_i(t)) \geq \sum_{i=0}^{n-1} -a^2(1-f_i(t))\cdot\frac{1}{1-f(0)} \geq \frac{-a^2}{1-f(0)}\sum_{i=0}^{n-1}(1-f_i(0)) \sim o(n).$$
We observe that
$$\begin{aligned}\sum_{i=0}^{n-1}\delta(f_i(t)) &= a-\left(\frac{1}{1-f(t)}-\frac{1}{1-t}\right)+\ldots+a-\left(\frac{1}{1-f_n(t)}-\frac{1}{1-f_{n-1}(t)}\right)\\ &= na+\frac{1}{1-t}-\frac{1}{1-f_n(t)}.\end{aligned}$$
Thus,
$$a-\frac{1}{n}\left(\frac{1}{1-f_n(t)}-\frac{1}{1-t}\right) \sim o(1).$$
□

3.2 The Conditioned Process

In order to avoid the inevitable extinction of the critical process, we may be interested in conditioning the process on non-extinction. Let $\{Z_n \mid Z_n > 0\}$ denote the process $\{Z_n\}$ under $\mathbb{P}_{\{Z_n>0\}}$. In the critical case, observe a beneficial relationship linking the expected value of the conditioned process at generation n with the probability of survival at that same generation. Namely,
$$1 = \mathbb{E}(Z_n) = \mathbb{P}(Z_n > 0)\cdot\mathbb{E}(Z_n \mid Z_n > 0) + \mathbb{P}(Z_n = 0)\cdot\mathbb{E}(Z_n \mid Z_n = 0).$$
Thus $\mathbb{E}(Z_n \mid Z_n > 0) = \mathbb{P}(Z_n > 0)^{-1}$, an inverse relationship that verifies the behavior of the conditioned process. Theorem 3.1.1 shows that the convergence $\mathbb{P}(Z_n > 0) \to 0$ is of the order n^{-1} and hence we have linear growth in the mean of the conditioned process. Cutting down the process by a factor of n yields the following result, from [3].

Theorem 3.2.1. *Let $m \doteq \sum_{j=1}^{\infty} jp_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2p_j - 1 < \infty$. Let $Z_0 = k < \infty$. Then as $n \to \infty$ and for $0 < u < \infty$,*
$$\mathbb{P}\big(Z_n > nu \mid Z_n > 0\big) \to e^{\frac{-2u}{\sigma^2}}.$$
Remark: Equivalently, $\lim_{n\to\infty}\mathbb{P}(\frac{Z_n}{n} > u \mid Z_n > 0) = e^{-\frac{2}{\sigma^2}u}$ and so $\{\frac{Z_n}{n} \mid Z_n > 0\}$ converges in distribution to Y where Y is an exponentially distributed random variable with parameter $\frac{2}{\sigma^2}$.

Proof. By Levy's Continuity Theorem, Theorem 2.1.2, a sequence of random variables X_n converges in distribution to X if φ_{X_n} converges in the pointwise sense to φ_X where φ_X is the characteristic function of the random variable X. Thus, we must look at the characteristic function for $\frac{Z_n}{n}$ under $\mathbb{P}_{\{Z_n>0\}}$.

$$\begin{aligned}
\varphi_{\frac{Z_n}{n}}(\theta) &= \mathbb{E}_{\mathbb{P}_{\{Z_n>0\}}}\Big((e^{i\frac{\theta}{n}})^{Z_n}\Big) \\
&= \sum_{k=1}^{\infty} e^{i\frac{\theta}{n}k}\mathbb{P}(Z_n = k \mid Z_n > 0) \\
&= \frac{\sum_{k=1}^{\infty} e^{i\frac{\theta}{n}k}\mathbb{P}(Z_n = k)}{\mathbb{P}(Z_n > 0)} \\
&= \frac{\sum_{k=0}^{\infty} e^{i\frac{\theta}{n}k}\mathbb{P}(Z_n = k) - \mathbb{P}(Z_n = 0)}{1-\mathbb{P}(Z_n = 0)} \\
&= \frac{f_n(e^{i\frac{\theta}{n}}) - f_n(0)}{1 - f_n(0)} \\
&= \frac{\big(f_n(e^{i\frac{\theta}{n}}) - 1\big) + \big(1 - f_n(0)\big)}{1 - f_n(0)} \\
&= 1-\left(\frac{1 - f_n(e^{i\frac{\theta}{n}})}{1 - f_n(0)}\right) \\
&= 1-\left(\frac{1}{1 - f_n(0)}\right)\left(\frac{1}{\frac{1}{1 - f_n(e^{i\frac{\theta}{n}})}}\right) \\
&= 1-\left(\frac{1}{n(1 - f_n(0))}\right)\left(\frac{1}{\frac{1}{n(1 - f_n(e^{i\frac{\theta}{n}}))} - \frac{1}{n(1 - e^{i\frac{\theta}{n}})} + \frac{1}{n(1 - e^{i\frac{\theta}{n}})}}\right)
\end{aligned}$$

Recall now the uniform convergence stated in (3.1). Because this convergence is uniform, we may consider $t_n = e^{i\frac{\theta}{n}}$ and letting $n \to \infty$ we obtain

$$\begin{aligned}
\varphi_{\frac{Z_n}{n}}(\theta) &\to 1 - \frac{1}{a}\left(\frac{1}{\frac{1}{a} + \lim_{n\to\infty} \frac{1}{n(1 - e^{i\frac{\theta}{n}})}}\right) \\
&= 1 - \frac{1}{a}\left(\frac{1}{\frac{1}{a} + \frac{1}{-i\theta}}\right), \\
&= 1 - \frac{1}{1+\frac{a}{\theta}i} \\
&= \frac{\frac{a}{\theta}i}{1 + \frac{a}{\theta}i} \\
&= \frac{1}{1 - \frac{\theta}{a}i}.
\end{aligned}$$

This is continuous at 0 as is required by Theorem 2.1.2 and is a well-known characteristic function. The density function of the exponential distribution with parameter λ is given by $\lambda e^{-\lambda x}\chi_{(0,\infty)}$. It is easy to see that

$$\int_{\mathbb{R}} e^{i\theta x}\lambda e^{-\lambda x}\chi_{(0,\infty)}\,dx = \frac{\lambda}{\lambda - i\theta} = \frac{1}{1-\frac{\theta}{\lambda}i}.$$

Thus, the characteristic function to which sequence of characteristic functions converges to is that of the exponential distribution with parameter $\lambda = a = \frac{2}{\sigma^2}$. So we have seen that

$$\mathbb{P}(Z_n \le un \mid Z_n > 0) \to F_{\exp}(u) \tag{3.2}$$

where $F_{\exp}(u)$ is the distribution function for the exponential random variable with parameter λ, thus proving the equivalent theorem statement.

□

Now we present the first original result that builds to our larger research goals in the sections to come.

Theorem 3.2.2. *Let $m \doteq \sum_{j=1}^{\infty} jp_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. Let $Z_0 = k < \infty$. Then as $n \to \infty$ and for $0 < u < \infty$,*

$$n\mathbb{P}(Z_n > nu) \to \frac{2}{\sigma^2}e^{-\frac{2}{\sigma^2}u}.$$

Remark: We know that with $m = 1$ the process goes extinct eventually with probability 1. Clearly, then $\lim_{n\to\infty}\mathbb{P}(Z_n > nu) = 0$, but the rate of convergence is clarified in this theorem. Note that in particular, this result tells us that $n\mathbb{P}(Z_n > n) \to \frac{2}{\sigma^2}e^{-\frac{2}{\sigma^2}}$ as $n \to \infty$.

Proof. From (3.2),

$$\frac{n\mathbb{P}(Z_n > nu)}{n\mathbb{P}(Z_n > 0)} = \mathbb{P}(Z_n > nu \mid Z_n > 0) \to 1 - F_{\exp}(u),$$

where

$$1 - F_{\exp}(u) = \int_u^{\infty} \lambda e^{-\lambda x}\,dx = e^{-\lambda u},$$

with $\lambda = \frac{2}{\sigma^2}$. By Theorem 3.1.1, $\lim_{n\to\infty} n\mathbb{P}(Z_n > 0) = \frac{2}{\sigma^2}$. Thus, as $n \to \infty$,

$$n\mathbb{P}(Z_n > nu) \to (1 - F_{\exp}(u)) \cdot \lim_{n\to\infty} n\mathbb{P}(Z_n > 0) = \frac{2}{\sigma^2}e^{-\frac{2}{\sigma^2}u}.$$

□

CHAPTER

4

ASYMPTOTIC BEHAVIOR OF THE TOTAL COALESCENCE TIME

In the following, we introduce the notion of coalescence. Let $Z^{(k)}_{n-k,i}$ denote the size of the $(n-k)$-th generation of a branching process initiated by individual i in the generation k. Then individual i in generation k is said to be a common ancestor of two individuals in generation $n \geq k$ if those individuals count towards the random variable $\{Z^{(k)}_{n-k,i}\}$.

Definition 4.0.1. *The total coalescence time is the generation of the most recent ancestor for any whole generation of the branching process. Denote by Y_n the random variable representing the total coalescence time for generation n.*

See [10] for some results pertaining to the coalescence time for branching processes. To find the value of the total coalescence time, trace back the lineage of each of the Z_n individuals in generation n until their distributions meet at one common ancestor. The generation number of this ancestor is the value of Y_n. Based on this definition, we note that for any $k \in \mathbb{N}$, the event $\{Y_n \geq k\}$ occurs if each of the Z_n individuals from generation n are offspring from a single individual in generation k, i.e. from generation $n-k$ of a branching process initiating with that single individual.
For ease of notation, let $I_m \doteq [1, \ldots, m]$. Then the event $\{Y_n \geq k\}$ occurs if and only if

$Z_{n-k,i}^{(k)} = 0$ for all but exactly one i, where i takes values in I_{Z_k}. This fact will be used as a starting point for the proofs of Theorems (4.1.1)-(4.1.5), (4.2.4) and (4.2.5). It must also be noted that the conditioning on the event $\{Z_0 = 1\}$ is suppressed but is always implied. It is shown when it clarifies the direction of the argument.

4.1 Conditioning on Inequalities

The first theorem given in this section from [1] is Theorem 4.1.1 and it was first proved by Zubkov. It provides an intuitive understanding between the distribution of the total coalescence time Y_n and that of Z_n asymptotically. Theorems 4.1.2-4.1.5, help to gain a more complete understanding of Y_n under multi-scale conditioning.

Cases where the coalescence time is bounded below by functions of order 1 and of order n are considered in our research to follow - reasonably so due to the nature of bounding below the coalescence time for generation n. The population size of the n-th generation will be considered in various cases as well, bounding it below by functions of various orders with respect to n. There are less restrictions on this conditioning, a result of the open-ended potential of the random variable Z_n, which is reflected in the results to come.

4.1.1 Case I

Theorem 4.1.1. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Then for* $0 < u < 1$,

$$\lim_{n\to\infty} \mathbb{P}(Y_n > nu \mid Z_n > 0) = 1 - u.$$

Remark: Equivalently, $\lim_{n\to\infty} \mathbb{P}(Y_n \leq nu \mid Z_n > 0) = u$. *Thus,* $\{\frac{Y_n}{n} \mid Z_n > 0\}$ *converges in distribution to* Y *where* Y *is uniformly distributed on* $[0,1]$.

Proof. We first consider $\mathbb{P}(Y_n \geq k \mid Z_n > 0)$, and then we will let $\frac{k}{n} \to u$ as $k, n \to \infty$. Letting $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$, we have by the remarks at the start of Chapter 4,

$$\mathbb{P}(Y_n \geq k \mid Z_n > 0) = \frac{\mathbb{P}(Z_{n-k,i}^{(k)} = 0 \text{ for all but one } i \in I_{Z_k},\ Z_k > 0)}{\mathbb{P}(Z_n > 0)} \tag{4.1}$$

Since $\sum_{m=1}^{\infty} f(m)\mathbb{P}(Z_k = m) = \mathbb{E}(f(Z_k))$,

$$\frac{\mathbb{P}(Z_{n-k,i}^{(k)} = 0 \text{ for all but one } i \in I_{Z_k},\ Z_k > 0)}{\mathbb{P}(Z_n > 0)} = \mathbb{E}\big(Z_k q_{n-k}^{Z_k - 1}(1 - q_{n-k}) \mid Z_k > 0\big)\frac{\mathbb{P}(Z_k > 0)}{\mathbb{P}(Z_n > 0)}. \tag{4.2}$$

In (4.2), we see the following terms appear. Z_k is the random variable representing the number of individuals in generation k, that is, the number of choices to make for which one of the individuals will "live on" to generation n. $q_{n-k} = \mathbb{P}(Z_{n-k} = 0)$ and so $q_{n-k}^{Z_k-1}$ gives the probability that each other individual in the k-th generation dies out by generation n. $(1-q_{n-k}) = \mathbb{P}(Z_{n-k} > 0)$, thus giving the probability of the one individual living to generation n. The factor $\mathbb{P}(Z_k > 0)$ is gained due to the conditioning on the event that $Z_k > 0$), which is a necessity in order to consider the events previously listed. Regrouping terms of the same order yields

$$\mathbb{P}(Y_n \geq k \mid Z_n > 0) = \mathbb{E}\left(\frac{Z_k}{k} q_{n-k}^{Z_k-1}(n-k)(1-q_{n-k}) \mid Z_k > 0\right)\frac{\mathbb{P}(Z_k > 0)}{\mathbb{P}(Z_n > 0)}\frac{k}{n-k}.$$

Under the given conditions, we may apply previously stated theorems to obtain the following results as $n, k \to \infty$ and $\frac{k}{n} \to u$.

Theorem 3.1.1 allows for considering the asymptotic behavior of $\mathbb{P}(Z_n > 0)$ as $n \to \infty$ and more specifically its rate of convergence to 0. Noting that as $n \to \infty$, $k \to \infty$ as well, we have

$$\begin{aligned}
\lim_{n\to\infty} \frac{\mathbb{P}(Z_k > 0)}{\mathbb{P}(Z_n > 0)} &= \lim_{n\to\infty} \frac{k\mathbb{P}(Z_k > 0)n}{n\mathbb{P}(Z_n > 0)k} \\
&= \frac{2/\sigma^2}{2/\sigma^2}\frac{1}{u} \qquad (4.3)\\
&= \frac{1}{u}.
\end{aligned}$$

Next, it is clear to see that

$$\begin{aligned}
\lim_{n\to\infty} \frac{k}{n-k} &= \lim_{n\to\infty} \frac{\frac{k}{n}}{1-\frac{k}{n}} \qquad (4.4)\\
&= \frac{u}{1-u}.
\end{aligned}$$

Now we must consider the conditional expectation $\mathbb{E}\left(\frac{Z_k}{k} q_{n-k}^{Z_k-1}(n-k)(1-q_{n-k}) \mid Z_k > 0\right)$. We see that what we have is the expected value of a convergent functional of a convergent sequence of random variables. Let $X_k \doteq \frac{Z_k}{k}$. By Theorem 3.2.1, $\{X_k \mid Z_k > 0\}$ asymptotically converges in distribution to Y where Y is an exponentially distributed random variable with parameter $\frac{2}{\sigma^2}$. We then know that the conditional expectation will converge to the expectation of the functional defined at the limit of the sequence of random variables in question, when computed under $\mathbb{P}_{\{Z_k>0\}}$.

We turn next to the limiting behavior of $q_{n-k}^{Z_k-1}$ as $n, k \to \infty$. In this case, we must take note that q_{n-k} is itself a functional, here acting on a sequence of random variables. As such, q_{n-k} is changing as n and k change, as is the sequence of random variables Z_k. In order to consider the asymptotic behavior, we may note that by nature of q_{n-k} as a probability function, it is uniformly bounded because its value will always fall between 0 and 1. Because of this, even $(q_{n-k})^M$ will be uniformly bounded above by 1 for any real number M. In our case, M is the value of a random variable, so this uniform bounding is key. As a result, we may use the Lebesgue Dominated Convergence Theorem to compute the limit in question. First we expand the following:

$$q_{n-k}^{Z_k-1} = (q_{n-k}^{n-k})^{\frac{Z_k-1}{n-k}}. \tag{4.5}$$

Now,

$$\begin{aligned} q_{n-k}^{n-k} &= \left(1-\frac{(n-k)(1-q_{n-k})}{n-k}\right)^{n-k} \\ &= \left(1+\frac{1}{n-k}(-(n-k)(1-q_{n-k}))\right)^{\frac{1}{1/(n-k)}}. \end{aligned} \tag{4.6}$$

We also know that

$$\begin{aligned} \lim_{n\to\infty}(n-k)(1-q_{n-k}) &= \lim_{n\to\infty}(n-k)\mathbb{P}(Z_{n-k} > 0) \\ &= \frac{2}{\sigma^2}, \end{aligned} \tag{4.7}$$

as a direct application of Theorem 3.1.1 and since $n-k \to \infty$ under the asymptotic behavior considered. From L'Hopital's Rule, it is clear that for $y \in \mathbb{R}$,

$$\begin{aligned} \lim_{x\to 0^+}(1+yx)^{1/x} &= \exp\big(\lim_{x\to 0^+}\frac{1}{x}\ln(1+yx)\big) \\ &= \exp\big(\lim_{x\to 0^+}\frac{y}{1+x}\big) \\ &= e^y. \end{aligned} \tag{4.8}$$

So, q_{n-k}^{n-k} is uniformly bounded and converges in distribution to $e^{-\frac{2}{\sigma^2}}$ under $\mathbb{P}_{\{Z_k>0\}}$. We now need to determine the asymptotic behavior of the random functional $\frac{Z_k-1}{n-k}$ under $\mathbb{P}_{\{Z_k>0\}}$. First,

$$\frac{Z_k-1}{n-k} = \left(\frac{Z_k-1}{k}\right)\left(\frac{k/n}{1-k/n}\right). \tag{4.9}$$

We may observe the asymptotic behavior of (4.9) as $n\to\infty$, $k\to\infty$, and $\frac{k}{n}\to u$. First, by

Theorem 3.2.1

$$\begin{aligned}\lim_{n\to\infty}\frac{Z_k-1}{k} &= \lim_{n\to\infty}\left(\frac{Z_k}{k}-\frac{1}{k}\right)\\ &= \frac{\sigma^2}{2}\eta,\end{aligned}\tag{4.10}$$

in distribution under $\mathbb{P}_{\{Z_k>0\}}$ where η is an exponentially distributed random variable with parameter 1. (4.4) has previously considered the remaining limit, verifying that as $n,k\to\infty$,

$$\frac{Z_k-1}{n-k}\to\frac{\sigma^2}{2}\eta\frac{u}{1-u}.\tag{4.11}$$

By way of combining (4.5)-(4.11), we arrive at the asymptotic behavior of the functional in question as $n,k\to\infty$ and $\frac{n}{k}\to u$, namely that the Lebesgue Dominated Convergence Theorem yields

$$\begin{aligned}q_{n-k}^{Z_k-1} &\to \exp\!\left(\frac{-2}{\sigma^2}\cdot\frac{\sigma^2}{2}\cdot\eta\cdot\frac{u}{1-u}\right)\\ &= \exp\!\left(-\frac{u}{1-u}\cdot\eta\right),\end{aligned}\tag{4.12}$$

in distribution under $\mathbb{P}_{\{Z_k>0\}}$ where η is an exponentially distributed random variable with parameter 1. Therefore, taking the expected value of the convergent sequence of functionals evaluated at a sequence random variables (X_k), converging under $\mathbb{P}_{\{Z_k>0\}}$, will yield a finite result. In particular, Theorems 3.2.1 and 2.1.1, and (4.7) and (4.12) yield that as $n,k\to\infty$ and $\frac{n}{k}\to u$,

$$\begin{aligned}\mathbb{E}\left(\frac{Z_k}{k}q_{n-k}^{Z_k-1}(n-k)(1-q_{n-k})\mid Z_k>0\right) &\to \mathbb{E}\left(\eta\exp\!\left(-\frac{u}{1-u}\eta\right)\right)\\ &= \int_0^\infty x\exp\!\left(-\frac{u}{1-u}x\right)e^{-x}\,dx\\ &= \frac{1}{\left(1+\frac{u}{1-u}\right)^2}.\end{aligned}\tag{4.13}$$

This follows from the convergence of $\frac{Z_k}{k}$ in the distributional sense under the conditional probability measure $\mathbb{P}_{\{Z_k>0\}}$. Above, we use the fact that for any $\theta>0$,

$$\int_0^\infty xe^{-\theta x}e^{-x}dx = \frac{1}{(1+\theta)^2}.\tag{4.14}$$

Therefore combining (4.13), (4.3), and (4.4),

$$\lim_{n\to\infty} \mathbb{P}(Y_n \geq k \mid Z_n > 0) = \frac{1}{\left(1+\frac{u}{1-u}\right)^2}\frac{1}{1-u}$$
$$= 1-u.$$

□

4.1.2 Case II

Theorem 4.1.2. *Let $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. Consider the asymptotic behavior of the conditional probability,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n))$$

where $g(n)$ and $h(n)$ are functions of n. Suppose that $\lim_{n\to\infty} \frac{g(n)}{n} = c_1$ and $\lim_{n\to\infty} \frac{h(n)}{n} = c_2$ where $c_1 \in (0,1)$, $c_2 > 0$. Then,

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n)) = (1-c_1)e^{\frac{-2c_1 c_2}{\sigma^2(1-c_1)}}.$$

Proof. We first consider $\mathbb{P}(Y_n > k_1 \mid Z_n > k_2)$, and we will take $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n} \to c_2$ for $c_1 \in (0,1)$, $c_2 > 0$ as $n \to \infty$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\big)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}.$$

The term $(1-q_{n-k})$ in Theorem 4.1.1 represented the probability that the one individual from generation k would still be alive at generation n. Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes at least k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same order yields

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \frac{k_1}{n-k_1}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) \mid Z_{k_1} > 0\right)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}$$
$$= \mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2)\frac{k_1}{n-k_1}.$$

Now under the given conditions, we may apply previously stated theorems to obtain the following results as $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n} \to c_2$ as $n \to \infty$.

First consider

$$\lim_{n\to\infty} \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}.$$

Multiplying by one and applying Theorems 3.1.1 and 3.2.2 to the numerator and denominator respectively, we have

$$\begin{aligned}\lim_{n\to\infty} \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} &= \lim_{n\to\infty} \frac{k_1 \mathbb{P}(Z_{k_1} > 0) n}{n \mathbb{P}(Z_n > k_2) k_1} \\ &= e^{2c_2/\sigma^2} \cdot \frac{1}{c_1}.\end{aligned} \tag{4.15}$$

Next, from (4.4),

$$\lim_{n\to\infty} \frac{k_1}{n-k_1} = \frac{c_1}{1-c_1}. \tag{4.16}$$

Now, we look at

$$(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) = (n-k_1)\mathbb{P}\Big(Z_{n-k_1} > \frac{k_2}{n-k_1}(n-k_1)\Big).$$

By applying Theorem 3.2.2 and noting that as $n \to \infty$, $n - k_1 \to \infty$, we obtain

$$\lim_{n\to\infty} (n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) = \frac{2}{\sigma^2} e^{\frac{-2c_2}{\sigma^2(1-c_1)}}, \tag{4.17}$$

in the distributional sense. It remains to consider the conditional expectation

$$\mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\Big).$$

This can be computed in much the same way as was done in Theorem 4.1.1. Again, by Theorem 3.2.1, $\frac{Z_{k_1}}{k_1}$ under $\mathbb{P}_{\{Z_{k_1} > 0\}}$ converges in distribution to Y where Y is an exponentially distributed random variable with parameter $\frac{2}{\sigma^2}$. As shown in (4.12), $q_{n-k_1}^{Z_{k_1}-1}$ under $\mathbb{P}_{\{Z_{k_1} > 0\}}$ converges in distribution to $e^{-\frac{c_1}{1-c_1}\eta}$, where η is an exponentially distributed random variable with parameter 1.

Putting these together, as $n, k_1 \to \infty$ and $\frac{k_1}{n} \to c_1$,

$$\begin{aligned}\mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\Big) &\to \mathbb{E}\Big(\frac{\sigma^2}{2}\eta \exp\big(-\frac{c_1}{1-c_1}\eta\big)\Big) \\ &= \frac{\sigma^2}{2}\int_0^\infty x \exp\big(-\frac{c_1}{1-c_1}x\big) e^{-x}\, dx \\ &= \frac{\sigma^2}{2} \frac{1}{\big(1+\frac{c_1}{1-c_1}\big)^2}.\end{aligned} \tag{4.18}$$

This follows from the convergence of $\frac{Z_{k_1}}{k_1}$ in the distributional sense under the conditional probability measure $\mathbb{P}_{\{Z_{k_1}>0\}}$. Again, we used (4.14) to compute the integral. Combining (4.15) -(4.18),

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(Y_n > k_1 \mid Z_n \geq k_2) &= \frac{\sigma^2}{2}\frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2}\frac{2}{\sigma^2}e^{\frac{-2c_2}{\sigma^2(1-c_1)}}e^{2c_2/\sigma^2}\cdot\frac{1}{c_1}\frac{c_1}{1-c_1}\\ &= (1-c_1)e^{\frac{-2c_2}{\sigma^2(1-c_1)}}e^{2c_2/\sigma^2}\\ &= (1-c_1)e^{\frac{-2c_1c_2}{\sigma^2(1-c_1)}}.\end{aligned}$$

□

4.1.3 Case III

Theorem 4.1.3. *Let $m \doteq \sum_{j=1}^{\infty} jp_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2p_j - 1 < \infty$. Consider the asymptotic behavior of the conditional probability,*

$$\lim_{n\to\infty}\mathbb{P}(Y_n > g(n) \mid Z_n > h(n))$$

where $g(n)$ and $h(n)$ are functions of n. Suppose that $\lim_{n\to\infty} g(n) = c_1$ and $\lim_{n\to\infty} h(n) = c_2$ where $c_1,\ c_2 \in \mathbb{N}$. Then,

$$\lim_{n\to\infty}\mathbb{P}(Y_n > g(n) \mid Z_n > h(n)) = 1.$$

Remark: This result confirms that when $h(n) > 0$, we achieve an the same resulting limit for $g(n)$ of order 1. In fact, it can be shown that when $h(n)$ is of any order less than or equal to n, this same limit holds for $g(n)$ of order 1.

Proof. We first consider $\mathbb{P}(Y_n > k_1 \mid Z_n > k_2)$, and then we will let $k_1(n) \to c_1$ and $k_2(n) \to c_2$ as $n \to \infty$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\Big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1}\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}.$$

Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes at least k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same

order yields

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\Big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{(n-k_1)\mathbb{P}(Z_n > k_2)}$$
$$= \mathbb{E}\Big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{(n-k_1)\mathbb{P}(Z_n > k_2)}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2).$$

Note that in the line above, $\mathbb{P}(Z_{k_1} > 0)$ is asymptotically constant. However, since Z_{k_1} is still not a deterministic term, we only consider that $\mathbb{E}(g(Z_{k_1}) \mid Z_{k_1} > 0)$ becomes constant asymptotically. This slight change from the previous cases drastically alters the way the asymptotic behavior of the remaining terms must be computed. Under the given conditions, we may apply previously stated theorems to obtain the following results as $n \to \infty$ and consequently, $k_1 \to c_1$ and $k_2 \to c_2$.

First let us consider the asymptotic behavior of $(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2)$ by looking at $r\mathbb{P}(Z_r > k)$ where k is a constant and $r \doteq n - k_1 \to \infty$ as $n \to \infty$. The definition of conditional probability gives

$$\frac{r\mathbb{P}(Z_r > k)}{r\mathbb{P}(Z_r > 0)} = \mathbb{P}(Z_r > k \mid Z_r > 0).$$

As such, it will be useful to simply compute the asymptotic behavior of

$$\mathbb{P}(Z_r > k \mid Z_r > 0) \cdot r\mathbb{P}(Z_r > 0).$$

We may compute the two parts separately since they will both have finite limits. By Theorem 3.2.1,

$$\begin{aligned}\lim_{r\to\infty} \mathbb{P}(Z_r > k \mid Z_r > 0) &= \lim_{r\to\infty} \mathbb{P}(\frac{Z_r}{r} > \frac{k}{r} \mid Z_r > 0) \\ &= 1,\end{aligned} \tag{4.19}$$

since $\frac{Z_r}{r} \xrightarrow{d} Y$ under $\mathbb{P}_{\{Z_r > 0\}}$, where Y has exponential distribution with parameter $\frac{2}{\sigma^2}$. Then by (4.19) and Theorem 3.1.1, and due to the continuity of conditional probability measure,

$$\begin{aligned}\lim_{r\to\infty} r\mathbb{P}(Z_r > k) &= \lim_{r\to\infty} \mathbb{P}\Big(\frac{Z_r}{r} > \frac{k}{r} \mid Z_r > 0\Big) \cdot r\mathbb{P}(Z_r > 0) \\ &= 1 \cdot \frac{2}{\sigma^2} \\ &= \frac{2}{\sigma^2}.\end{aligned} \tag{4.20}$$

So, in (4.20) we may replace r by $n-k_1$ once again to see that

$$\lim_{n\to\infty}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) = \frac{2}{\sigma^2}. \tag{4.21}$$

As a result, we are also able to compute the asymptotic behavior of the following:

$$\begin{aligned}
\lim_{n\to\infty}(n-k_1)\mathbb{P}(Z_n > k_2) &= \lim_{n\to\infty}\frac{n-k_1}{n}\; n\mathbb{P}(Z_n > k_2) \\
&= 1\cdot\frac{2}{\sigma^2} \\
&= \frac{2}{\sigma^2}.
\end{aligned} \tag{4.22}$$

Next, we consider the asymptotic behavior of $q_{n-k_1}^{Z_{k_1}-1}$ in order to compute the desired conditional expectation. We may use algebra to rewrite the following:

$$\begin{aligned}
q_{n-k_1}^{Z_{k_1}-1} &= \exp\big[(Z_{k_1}-1)\ln(q_{n-k_1})\big] \\
&= \exp\big[(Z_{k_1}-1)\ln\big(1-(1-q_{n-k_1})\big)\big] \\
&= \exp\left[(Z_{k_1}-1)(1-q_{n-k_1})\frac{\ln\big(1-(1-q_{n-k_1})\big)}{1-q_{n-k_1}}\right].
\end{aligned}$$

This algebra is particularly advantageous since in the critical case considered, $q_r \to 1$ as $r\to\infty$ by definition and so $1-q_{n-k_1}\to 0$ as $n\to\infty$. From $\lim_{u\to 0}\frac{\ln(1-u)}{u} = -1$, we obtain

$$\begin{aligned}
\lim_{n\to\infty} q_{n-k_1}^{Z_{k_1}-1} &= \lim_{n\to\infty}\exp\left[(Z_{k_1}-1)(1-q_{n-k_1})\frac{\ln\big(1-(1-q_{n-k_1})\big)}{1-q_{n-k_1}}\right] \\
&= 1,
\end{aligned} \tag{4.23}$$

in distribution, under $\mathbb{P}_{\{Z_n>0\}}$. It remains to consider the asymptotic behavior of the conditional expectation as well as the remaining term outside of the expectation:

$$\mathbb{E}\left(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right)\mathbb{P}(Z_{k_1} > 0).$$

As $n\to\infty$, $Z_{k_1}\to Z_{c_1}$, the population size at generation c_1, in the pointwise sense. So it is left to look at $\mathbb{E}\big(Z_{c_1}\cdot 1 \mid Z_{c_1} > 0\big)\mathbb{P}(Z_{c_1} > 0)$. Let the event $A \doteq \{Z_{c_1} > 0\}$. Recall that $\mathbb{P}_A(B) = \mathbb{P}(B\mid A)$ for all events $A\in\mathscr{A}$. Now since $d\mathbb{P}_A$ is absolutely continuous with respect to $d\mathbb{P}$, the Radon-Nikodym derivative $\frac{d\mathbb{P}_A}{d\mathbb{P}}$ exists and it is $\frac{d\mathbb{P}_A}{d\mathbb{P}} = \frac{1}{\mathbb{P}(A)}\chi_A$. Thus

$$\begin{aligned}\mathbb{E}(Z_{c_1} \mid A)\mathbb{P}(A) &= \frac{1}{\mathbb{P}(A)}E(Z_{c_1} \cdot \chi_A)\mathbb{P}(A) \\ &= \mathbb{E}(Z_{c_1}) \\ &= 1,\end{aligned} \tag{4.24}$$

since we are dealing with the critical case for the branching process. Combining the results from (4.21), (4.22), and (4.24), we have

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n > k_2) &= 1 \cdot \frac{1}{2/\sigma^2}\frac{2}{\sigma^2} \\ &= 1.\end{aligned}$$

□

4.1.4 Case IV

Theorem 4.1.4. *Let* $m \doteq \sum_{j=1}^{\infty} jp_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Consider the asymptotic behavior of the conditional probability,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n))$$

where $g(n)$ *and* $h(n)$ *are functions of* n. *Suppose that* $\lim_{n\to\infty} \frac{g(n)}{n} = c_1$ *and* $\lim_{n\to\infty} h(n)n^{\alpha} = c_2$ *where* c_1, $\alpha \in (0,1)$ *and* $c_2 > 0$. *Then,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n)) = 1 - c_1.$$

Remark: Recall the result in Theorem 4.1.1 and compare it to the result above. This theorem clarifies that for $h(n)$ *of any order strictly less than* n*, we achieve the same result in this case, conditioning on* $\{Z_n > h(n)\}$*, as with conditioning on non-extinction where* $h(n) \equiv 0$.

Proof. Now we consider $\mathbb{P}(Y_n > k_1 \mid Z_n > k_2)$, assuming that $\frac{k_1}{n} \to c_1$ and $k_2 n^{\alpha} \to c_2$ as $n \to \infty$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\Big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}.$$

Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes at least k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same order yields

$$\begin{aligned}
\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) &= \mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \frac{k_1}{n-k_1}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2) \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \\
&= \mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\Big)\frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \cdot (n-k_1)\mathbb{P}(Z_{n-k_1} > k_2)\frac{k_1}{n-k_1}.
\end{aligned}$$

Under the given conditions, we may apply previously stated theorems to obtain the following results as $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n^\alpha} \to c_2$ as $n \to \infty$.

First consider $\lim_{n\to\infty} \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}$. Scaling appropriately and applying Theorems 3.1.1 and 3.2.2 respectively, we obtain

$$\begin{aligned}
\lim_{n\to\infty} \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} &= \lim_{n\to\infty} \frac{k_1\mathbb{P}(Z_{k_1} > 0)n}{n\mathbb{P}(Z_n > k_2)k_1} \\
&= \frac{2/\sigma^2}{\frac{2}{\sigma^2}\exp\big(-\frac{2}{\sigma^2}\lim_{n\to\infty} n^{\alpha-1}c_2\big)} \cdot \frac{1}{c_1} \\
&= \frac{1}{c_1}\exp\Big(\frac{2}{\sigma^2}\lim_{n\to\infty} n^{\alpha-1}c_2\Big) \\
&= \frac{1}{c_1}.
\end{aligned} \tag{4.25}$$

In the equation above, we compute the asymptotic behavior of $n\mathbb{P}(Z_n > k_2)$ via Theorem 3.2.2 and taking advantage of the continuity of the exponential function. Next from (4.4),

$$\lim_{n\to\infty} \frac{k_1}{n-k_1} = \frac{c_1}{1-c_1}. \tag{4.26}$$

Now, by Theorem 3.2.2 and since $\lim_{n\to\infty} \frac{k_2}{n-k_1} = 0$ in the present case,

$$\begin{aligned}
\lim_{n\to\infty}(n-k_1)\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) &= \lim_{n\to\infty}(n-k_1)\mathbb{P}\Big(Z_{n-k_1} > \frac{k_2}{n-k_1}(n-k_1)\Big) \\
&= \frac{2}{\sigma^2}.
\end{aligned} \tag{4.27}$$

It remains to consider the conditional expectation $\mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\Big)$. Because of the limiting nature of k_1 being identical to that in Theorem 4.1.2, the same proof technique shown in (4.18) is employed here to verify that as $n \to \infty$ and concurrently $\frac{k_1}{n} \to c_1$ and

$\frac{k_2}{n^\alpha} \longrightarrow c_2$,

$$\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right) \to \frac{\sigma^2}{2} \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2}. \tag{4.28}$$

This follows from the convergence of $\frac{Z_{k_1}}{k_1}$ in the distributional sense under the conditional probability measure $\mathbb{P}_{\{Z_{k_1}>0\}}$. Combining the results from (4.25)-(4.28),

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n > k_2) &= \frac{\sigma^2}{2} \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2} \frac{1}{c_1} \frac{2}{\sigma^2} \frac{c_1}{1-c_1} \\
&= \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2} \frac{1}{1-c_1} \\
&= 1-c_1.
\end{aligned}$$

□

4.1.5 Case V

Theorem 4.1.5. *Let* $m \doteq \sum_{j=1}^\infty j p_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^\infty j^2 p_j - 1 < \infty$. *Consider the asymptotic behavior of the conditional probability,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n))$$

where $g(n)$ *and* $h(n)$ *are functions of* n. *Suppose that* $\lim_{n\to\infty} \frac{g(n)}{n} = c_1$ *and* $\lim_{n\to\infty} \frac{h(n)}{n^\alpha} = c_2$ *where* $c_1 \in (0,1)$, $c_2 > 0$, *and* $\alpha \in (1,\infty)$. *Then,*

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n > h(n)) = \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2}.$$

Proof. We first consider $\mathbb{P}(Y_n > k_1 \mid Z_n > k_2)$, and then we will let $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n^\alpha} \to c_2$ for $\alpha > 1$ as $k_1, k_2, n \to \infty$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) = \mathbb{E}\left(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\right) \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)}.$$

Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes at least k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same

order yields

$$\begin{aligned}\mathbb{P}(Y_n > k_1 \mid Z_n > k_2) &= \mathbb{E}\left(\frac{Z_{k_1}}{k_1} k_1 q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} > k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\right) \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \\ &= \mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right) \frac{k_1 \mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \cdot \mathbb{P}(Z_{n-k_1} > k_2).\end{aligned}$$

Under the given conditions, we may apply previously stated results to obtain the following results as $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n^\alpha} \to c_2$ as $n \to \infty$.
First, by Theorem 3.1.1, we see that

$$\lim_{n\to\infty} k_1 \mathbb{P}(Z_{k_1} > 0) = \frac{2}{\sigma^2}. \tag{4.29}$$

Now we must consider the asymptotic behavior of

$$\frac{\mathbb{P}(Z_{n-k_1} > k_2)}{\mathbb{P}(Z_n > k_2)}. \tag{4.30}$$

We may wish to first understand the asymptotic behavior of the random variable $\frac{Z_n}{n^\alpha}$ under the conditional measure $\mathbb{P}_{\{Z_n>0\}}$. For this, consider the Laplace transform of $\frac{Z_n}{n^\alpha}$ under $\mathbb{P}_{\{Z_n>0\}}$.

$$\begin{aligned}\mathbb{E}(e^{-\theta \frac{Z_n}{n}} \mid Z_n > 0) &= \mathbb{E}(e^{-\frac{\theta}{n^\alpha} Z_n} \mid Z_n > 0) \\ &= \frac{f_n(e^{-\frac{\theta}{n^\alpha}} - f_n(0)}{1 - f_n(0)} \\ &= 1 - \frac{1 - f_n(e^{-\frac{\theta}{n^\alpha}})}{1 - f_n(0)} \\ &= 1 - \frac{\left(n(1 - f_n(0))\right)^{-1}}{\left(n(1 - f_n(e^{-\frac{\theta}{n^\alpha}}))\right)^{-1}}.\end{aligned} \tag{4.31}$$

Let $a \doteq 2/\sigma^2$ and recall from Theorem 3.1.1 that $\lim_{n\to\infty} n\mathbb{P}(Z_n > 0) = a$. Further, (3.1) affirms that

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{1 - f_n(t)} - \frac{1}{1-t}\right) = \frac{1}{a},$$

where this convergence is uniform for $t \in [0,1)$. Thus,

$$\lim_{n\to\infty} \left(n(1 - f_n(e^{-\frac{\theta}{n^\alpha}}))\right)^{-1} = \frac{1}{a} + \lim_{n\to\infty} \left(n(1 - e^{-\frac{\theta}{n^\alpha}})\right)^{-1}, \tag{4.32}$$

which diverges to ∞ since $\alpha > 1$. Using the limit from (4.32) in (4.31), we see that

$$\lim_{n\to\infty} \mathbb{E}(e^{-\theta \frac{Z_n}{n}} \mid Z_n > 0) = 1.$$

In conclusion, the random variable $\frac{Z_n}{n^\alpha}$ converges in distribution to 0 under the probability measure $\mathbb{P}_{\{Z_n>0\}}$. Equivalently, both the numerator and denominator of (4.30) converge to 0, recalling that $k_1/n \to c_1 \in (0,1)$.

It will be shown in Theorem 4.2.3 that under the same conditions as those considered in the present case, $n^2\mathbb{P}(Z_n = k)$ converges asymptotically to some finite limit, for any $k \in \mathbb{N}$. Then, by nature of $\mathbb{P}(Z_n > k_2)$ as a summation over $k \in \mathbb{N}$ of $\mathbb{P}(Z_n = k)$ beginning with an arbitrarily large number k_2, it is clear that the asymptotic behavior of $\mathbb{P}(Z_n > k_2)$ is the same as that of the integral

$$\int_{k_2}^{\infty} \frac{1}{x^2}\, dx = \frac{1}{k_2}.$$

Thus,

$$\lim_{k_2\to\infty} k_2 \int_{k_2}^{\infty} \frac{1}{x^2}\, dx = b,$$

where b is a finite constant. Regardless of the precise value of this constant, we may compute the desired limit from (4.30) to see that

$$\lim_{n\to\infty} \frac{\mathbb{P}(Z_{n-k_1} > k_2)}{\mathbb{P}(Z_n > k_2)} = 1. \tag{4.33}$$

It remains to consider the conditional expectation $\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right)$. Because of the limiting nature of k_1 being identical to that in Theorem 4.1.2, the same proof technique shown in (4.18) is employed here to verify that as $n \to \infty$ and concurrently $\frac{k_1}{n} \to c_1$ and $\frac{k_2}{n^\alpha} \to c_2$,

$$\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right) \to \frac{\sigma^2}{2} \frac{1}{\left(1 + \frac{c_1}{1-c_1}\right)^2}. \tag{4.34}$$

This follows from the convergence of $\frac{Z_{k_1}}{k_1}$ in the distributional sense under the conditional probability measure $\mathbb{P}_{\{Z_{k_1}>0\}}$. Combining (4.29), (4.33), and (4.34),

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n > k_2) &= \frac{2}{\sigma^2} \cdot \frac{\sigma^2}{2} \cdot \frac{1}{\left(1 + \frac{c_1}{1-c_1}\right)^2} \cdot 1 \\ &= \frac{1}{\left(1 + \frac{c_1}{1-c_1}\right)^2}\end{aligned}$$

□

4.2 Conditioning on Fixed Population Size

The goal in this section is to consider the limits of the form

$$\lim_{n\to\infty} \mathbb{P}(Y_n \ > \ g(n) \,|\, Z_n \ = \ h(n)).$$

In such cases, we are assuming that *exact* knowledge of the asymptotic behavior of the population size is known, under proper scaling. The culmination of this section are Theorems 4.2.4 and 4.2.5, which give the resulting limiting behavior where $g(n)$ is assumed to have asymptotic behavior dictated by two distinct functions of n. In both cases, $h(n)$ is assumed to converge to a natural number.

In order to compute the resulting limits and determine the desired asymptotic behavior for the total coalescence time, we need to investigate the right order of convergence of $\mathbb{P}(Z_n = k)$ to 0. Surely, this quantity will converge to 0 asymptotically, since it has already been seen in Theorem 3.1.1 that $\mathbb{P}(Z_n > 0)$ decays at a rate in the order of n^{-1}. Since

$$\mathbb{P}(Z_n > 0) \ = \ \sum_{i=1}^{\infty} \mathbb{P}(Z_n = i),$$

it is intuitively clear that $\mathbb{P}(Z_n = k)$ should decay at an even faster rate. One potential method for obtaining the proper scaling for this limit is to utilize theory of the Yaglom limit and the notion of the quasi-stationary distribution. Recall that in the case of the conditioned process, we have already seen many results pertaining to the Yaglom limit

$$\lim_{n\to\infty} \mathbb{P}(Z_n = k \mid Z_n > 0)$$

in Section 2.4.1.

Note that

$$\mathbb{P}(Z_n = k) \ = \ \mathbb{P}(Z_n = k \mid Z_n > 0)\, \mathbb{P}(Z_n > 0).$$

If one could observe that $\mathbb{P}(Z_n = k \mid Z_n > 0)$ converges to some finite and nontrivial constant, then it would be clear that $\mathbb{P}(Z_n = k)$ converges to 0 in the same fashion as $\mathbb{P}(Z_n > 0)$, which is known (See Theorem 3.1.1). Unfortunately, Theorem 2.4.6 shows that in the critical case,

$$\lim_{n\to\infty} \mathbb{P}(Z_n = k \mid Z_n > 0) = 0.$$

More refined conditioning on $\mathbb{P}(Z_n = k)$ is required to go one step closer to a full understanding on the conditional convergence result of the critical branching process. In fact, Theorem 2.4.6 proves to be helpful towards the end of constructing a non-degenerate limit. To this end, we state the following lemmas found in [3], but with a proof included to illuminate reasoning and show the usefulness of the generating function in this context.

Lemma 4.2.1. *If $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 > 0$, and $p_0 < 1$, then for any $k \in \mathbb{N}$,*

$$\lim_{n\to\infty} \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0) = \sum_{k=1}^{\infty} \frac{\pi_k p_0^k}{\sum \pi_k p_0^k} \doteq \theta_k,$$

where $\theta_k \geq 0$, $\sum_{k=1}^{\infty} \theta_k = 1$.

Proof. We consider $\mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0)$, where $k \in \mathbb{N}$. The additional conditioning on the event $\{Z_{n+1} = 0\}$ yields the term p_0^k in the following, because each of the k individuals from generation n must die out at generation $n+1$. Recall that $\{p_y\}_{y\in\mathbb{N}}$ denotes the i.i.d. offspring distribution of each of the individuals. Seeking to express $\mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0)$ in terms of π_k, we note that

$$\begin{aligned}
\mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0) &= \frac{\mathbb{P}(Z_n = k,\ Z_n > 0,\ Z_{n+1} = 0)}{\mathbb{P}(Z_n > 0,\ Z_{n+1} = 0)} \\
&= \frac{\mathbb{P}^n(1,k) p_0^k}{\sum_{i=1}^{\infty} \mathbb{P}^n(1,i) p_0^i} \\
&= \frac{\frac{\mathbb{P}^n(1,k) p_0^k}{\mathbb{P}^n(1,1)}}{\sum_{i=1}^{\infty} \frac{\mathbb{P}^n(1,i)}{\mathbb{P}^n(1,1)} p_0^i}
\end{aligned}$$

Taking the limit of each term as $n \to \infty$,

$$\frac{\mathbb{P}^n(1,k)}{\mathbb{P}^n(1,1)} \longrightarrow \pi_k$$

by Lemma 2.4.4, thus proving the desired convergence to θ_k defined in the lemma statement. This limit is well-defined because of the additional term p_0^k within the sum; In Theorem 2.4.5, $\sum_{k=1}^{\infty} \pi_k$ was shown to diverge in the critical case, however this denominator is now converted into a power series in $p_0 \in [0,1)$ where the π_ks yield the coefficients. $\square$

Lemma 4.2.2. *Let $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. Then*

$$\lim_{n\to\infty} n^2 \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) = \frac{2}{\sigma^2}.$$

Proof. Consider that

$$\mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) \;=\; \mathbb{P}(Z_n > 0) - \mathbb{P}(Z_n > 0,\ Z_{n+1} > 0) \tag{4.35}$$

This implies that

$$\mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) \;=\; \mathbb{P}(Z_n > 0) - \mathbb{P}(Z_{n+1} > 0). \tag{4.36}$$

Now,

$$\begin{aligned} \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) &= (1 - f_n(0)) - (1 - f_{n+1}(0)) \\ &= f_{n+1}(0) - f_n(0). \end{aligned} \tag{4.37}$$

We will take advantage of a previously given result about the generating function $f_n(t)$ of Z_n. From (3.1) in the proof of Theorem 3.1.1,

$$\lim_{n\to\infty} \frac{1}{n}\left(\frac{1}{1-f_n(t)} - \frac{1}{1-t}\right) = \frac{\sigma^2}{2},$$

uniformly on $[0,1)$. In this lemma, we wish to consider

$$f_{n+1}(t) - f_n(t).$$

By definition of $f_n(t)$,

$$f_{n+1}(t) - f_n(t) = f(f_n(t)) - f_n(t).$$

It is known that as $n \to \infty$, $f_n(t) \to 1$ and further that 1 is a fixed point of $f(t)$. Together, this gives that as $n \to \infty$, $f(f_n(t)) - f_n(t) \to 0$. The rate of convergence is of importance here to obtain the proper scaling for the result. Consider the following limit:

$$\begin{aligned} \lim_{n\to\infty} \frac{f(f_n(t)) - f_n(t)}{(f_n(t)-1)^2} &= \lim_{x\to 1^-} \frac{f(x) - x}{(x-1)^2} \\ &= \lim_{x\to 1^-} \frac{f'(x) - 1}{2(x-1)} \\ &= \lim_{x\to 1^-} \frac{f''(x)}{2} \\ &= \frac{\sigma^2}{2}. \end{aligned} \tag{4.38}$$

As in Theorem 3.1.1, we are using the fact that the generating function for the offspring distribution, f, is assumed to have finite variance, denoted by σ^2. Because of this assumption and its correspondence to the $\mathscr{C}^2$-differentiability of f, the limit above is well-defined.

Thus,

$$\frac{\sigma^2}{2} = \lim_{n\to\infty} \frac{f(f_n(t)) - f_n(t)}{(f_n(t)-1)^2} = \lim_{n\to\infty} \frac{n^2(f_{n+1}(t) - f_n(t))}{n^2(f_n(t)-1)^2}.$$

Since the result above from Theorem 3.1.1 gives $\lim_{n\to\infty} n(f_n(t)-1) = \frac{2}{\sigma^2}$, the denominator converges to $\frac{4}{\sigma^4}$. Then it must be the case that the numerator converges to $\frac{2}{\sigma^2}$, i.e.

$$\lim_{n\to\infty} n^2(f_{n+1}(t) - f_n(t)) = \frac{2}{\sigma^2}, \quad t < 1. \tag{4.39}$$

Applying this result for the case $t = 0$ to (4.37),

$$\begin{aligned} \lim_{n\to\infty} n^2 \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) &= \lim_{n\to\infty} n^2\big(f_{n+1}(0) - f_n(0)\big) \\ &= \frac{2}{\sigma^2}. \end{aligned}$$

□

We take advantage of the previous lemmas to conclude the following result that is useful for determining the limit considered in Theorem 4.2.4, the culmination of this chapter.

Theorem 4.2.3. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $0 < p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Then for any* $k \in \mathbb{N}$,

$$\lim_{n\to\infty} n^2 \mathbb{P}(Z_n = k) = \frac{2\theta_k}{\sigma^2 p_0^k}.$$

Proof.

$$\mathbb{P}(Z_n = k,\ Z_n > 0,\ Z_{n+1} = 0) = \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0)\, \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0). \tag{4.40}$$

Then

$$\mathbb{P}(Z_{n+1} = 0 \mid Z_n = k)\, \mathbb{P}(Z_n = k) = \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0)\, \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0). \tag{4.41}$$

Since p_0^k gives the probability of the k individuals each going extinct in the next generation,

$$p_0^k \mathbb{P}(Z_n = k) = \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0)\, \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0). \tag{4.42}$$

Now by Lemma 4.2.1, $\lim_{n\to\infty} \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0) = \theta_k$ and further, by Lemma 4.2.2, $\lim_{n\to\infty} n^2 \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) = \frac{2}{\sigma^2}$. Thus,

$$\lim_{n\to\infty} n^2 \mathbb{P}(Z_n = k) = \lim_{n\to\infty} \frac{1}{p_0^k} \mathbb{P}(Z_n = k \mid Z_n > 0,\ Z_{n+1} = 0) n^2 \mathbb{P}(Z_n > 0,\ Z_{n+1} = 0) = \frac{2\theta_k}{\sigma^2 p_0^k}.$$

$\square$

The following theorem is the culmination of what has been done in this chapter, and gives new insight into the asymptotic behavior of the total coalescence time under more precise conditioning than those previously considered.

Theorem 4.2.4. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1, 0 < p_1 < 1,$ *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty.$

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n = h(n)) = 1,$$

where $\lim_{n\to\infty} \frac{g(n)}{n} = c_1$ *and* $\lim_{n\to\infty} h(n) = c_2$ *with* $c_1 \in (0,1),\ c_2 \in \mathbb{N}.$
Remark: This result confirms that the same limit holds when conditioning on $h(n)$ *of order* 1 *and considering* $g(n)$ *of any order strictly less than* n.

Proof. We first consider $\mathbb{P}(Y_n > k_1 \mid Z_n = k_2)$, and then we will let $\frac{k_1}{n} \to c_1$ as $n \to \infty$ and $\lim_{n\to\infty} k_2 = c_2$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n = k_2) = \mathbb{E}\big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\big) \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n = k_2)}.$$

Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes exactly k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same order yields

$$\begin{aligned}
\mathbb{P}(Y_n > k_1 \mid Z_n = k_2) &= \mathbb{E}\Big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\Big) \frac{k_1 \mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n = k_2)} \\
&= \mathbb{E}\bigg(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\bigg) \frac{\mathbb{P}(Z_{n-k_1} = k_2)}{\mathbb{P}(Z_n = k_2)} k_1 \mathbb{P}(Z_{k_1} > 0).
\end{aligned}$$

Under the given conditions, we may apply previously stated results to obtain the following results as $\frac{k_1}{n} \to c_1$ and $k_2 \to c_2$ as $n \to \infty$.
First we consider the asymptotic behavior of the piece not yet handled by previous cases, that is,

$$\frac{\mathbb{P}(Z_{n-k_1} = k_2)}{\mathbb{P}(Z_n = k_2)}.$$

Let $r \doteq n - k_1$ and note that r is then of the order n and behaves as $(1-c_1)n$ asymptotically. By Theorem 4.2.3, $n^2 \mathbb{P}(Z_n = k_2) \to \frac{2\theta_{k_2}}{\sigma^2 p_0^{k_2}}$ as $n \to \infty$ and $k_2 \to c_2$. Therefore, as $n \to \infty$ and concurrently $\frac{k_1}{n} \to c_1$, and $k_2 \to c_2$,

$$\begin{aligned}
\frac{\mathbb{P}(Z_{n-k_1} = k_2)}{\mathbb{P}(Z_n = k_2)} &= \frac{r^2 \mathbb{P}(Z_r = k_2)}{n^2 \mathbb{P}(Z_n = k_2)} \frac{n^2}{r^2} \\
&= \frac{n^2}{r^2} \frac{\frac{2\theta_{k_2}}{\sigma^2 p_0^{k_2}}}{\frac{2\theta_{k_2}}{\sigma^2 p_0^{k_2}}} \\
&= \frac{n^2}{(n-k_1)^2} \\
&\to \frac{1}{(1-c_1)^2}.
\end{aligned} \tag{4.43}$$

Next, from Theorem 3.1.1 and since as $n \to \infty$, $k_1 \to \infty$,

$$\lim_{n\to\infty} k_1 \mathbb{P}(Z_{k_1} > 0) = \frac{2}{\sigma^2}. \tag{4.44}$$

It remains to consider the conditional expectation $\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right)$. Let us recall the corresponding part from Theorem 4.1.2 to verify that as $n \to \infty$ and concurrently $\frac{k_1}{n} \to c_1$, and $k_2 \to c_2$,

$$\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\right) \to \frac{\sigma^2}{2} \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2}. \tag{4.45}$$

Here we notice that this follows from the convergence of $\frac{Z_{k_1}}{k_1}$ in the distributional sense under the conditional probability measure $\mathbb{P}_{\{Z_{k_1}>0\}}$. Then combining (4.43)-(4.45),

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n = k_2) &= \frac{\sigma^2}{2} \frac{1}{\left(1+\frac{c_1}{1-c_1}\right)^2} \frac{1}{(1-c_1)^2} \frac{2}{\sigma^2} \\
&= 1.
\end{aligned}$$

□

The result of Theorem 4.2.4 seems a bit disappointing at first glance, but it gives some interesting new insight into the coalescence behavior in the critical case. With k_2 of order 1 and $\frac{k_1}{n} \to c_1 \in (0,1)$ as $n \to \infty$,

$$\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n = k_2) = 1,$$

regardless of whether c_1 is close to 0, close to 1, or anywhere in between. When c is arbitrarily close to 1, this theorem narrows the region of coalescence to a very small range of possibility, thus yielding a particularly powerful result. The resulting limiting probability in Theorem 4.2.4 being equal to one indicates that we are not putting a strong enough constraint on the asymptotic order of $g(n)$. We need to refine the order of the function $g(n)$ to obtain a result that dives deeper into this phenomenon.

Consider, then, the timeline for the branching process between generations $c_1 n$ and n. In the case described above, the n-th generation will coalesce in this interval with probability one, but perhaps more can be said by further decomposing the behavior of k_1. First, consider letting $\frac{k_1}{n-n^a} \to 1$ where $a \in (0,1)$ so that the order of k_1 is comparable to n, though strictly smaller than n, in opposition to the considerations in Theorem 4.2.4. Note that for n large enough, $n-n^a$ will lie in-between cn and n for any $c \in (0,1)$. See in the following that $\frac{n-n^a}{n} \to 1$ asymptotically.

$$\begin{aligned} \lim_{n\to\infty} \frac{n-n^a}{n} &= \lim_{n\to\infty} 1-n^{a-1} \\ &= 1, \end{aligned} \tag{4.46}$$

since $a \in (0,1)$. In the following theorem, this research presents the resulting limit under these considerations.

Theorem 4.2.5. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1, 0 < p_1 < 1,$ *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty.$

$$\lim_{n\to\infty} \mathbb{P}(Y_n > g(n) \mid Z_n = h(n)) = 0,$$

where $g(n) = n - n^a$ *with* $a \in (0,1)$*, and* $\lim_{n\to\infty} h(n) = c_1$ *with* $c_1 \in \mathbb{N}$.
Remark: Based on previous discussion, we may note that $\lim_{n\to\infty} \frac{g(n)}{n} = 1$.

Proof. We first consider $\mathbb{P}(Y_n > k_1 \mid Z_n = k_2)$, and then we will let $\frac{k_1}{n-n^a} \to 1$ as $n \to \infty$ and $\lim_{n\to\infty} k_2 = c_1 > 0$. As in Theorem 4.1.1, let $q_n \doteq \mathbb{P}(Z_n = 0 \mid Z_0 = 1)$. As a direct extension of (4.1) and (4.2), we have

$$\mathbb{P}(Y_n > k_1 \mid Z_n = k_2) = \mathbb{E}\big(Z_{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\big) \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n = k_2)}.$$

Based on our adjusted conditioning, we now require that this one individual not only live on, but that it also contributes exactly k_2 offspring to generation n. Thus, the term $\mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1)$ above is necessary in this case. Regrouping the terms of the same

order yields

$$\begin{aligned}
\mathbb{P}(Y_n > k_1 \mid Z_n = k_2) &= \mathbb{E}\big(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mathbb{P}(Z_{n-k_1} = k_2 \mid Z_0 = 1) \mid Z_{k_1} > 0\big) \frac{k_1 \mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n = k_2)} \\
&= \mathbb{E}\bigg(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} \mid Z_{k_1} > 0\bigg) \frac{\mathbb{P}(Z_{n-k_1} = k_2)}{\mathbb{P}(Z_n = k_2)} k_1 \mathbb{P}(Z_{k_1} > 0) \\
&= \mathbb{E}\bigg(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} n^{2(1-a)} \mid Z_{k_1} > 0\bigg) \frac{n^{2a} \mathbb{P}(Z_{n-k_1} = k_2)}{n^2 \mathbb{P}(Z_n = k_2)} k_1 \mathbb{P}(Z_{k_1} > 0).
\end{aligned}$$

Note that $n - k_1$ behaves asymptotically like n^a when $a \in (0,1)$. Under the given conditions, we may apply previously stated results to obtain the following results as $\frac{k_1}{n-n^a} = 1$ and $k_2 \to c_1$ as $n \to \infty$.

First, note that since k_1 behaves asymptotically like $n - n^a$,

$$\begin{aligned}
\lim_{n\to\infty} k_1 &= \lim_{n\to\infty} n - n^a \\
&= \infty
\end{aligned}$$

So from Theorem 3.1.1,

$$\lim_{n\to\infty} k_1 \mathbb{P}(Z_{k_1} > 0) = \frac{2}{\sigma^2}. \tag{4.47}$$

Next, we consider the asymptotic behavior of the ratio

$$\frac{n^{2a} \mathbb{P}(Z_{n-k_1} = k_2)}{n^2 \mathbb{P}(Z_n = k_2)}.$$

By Theorem 4.2.3, $m^2 \mathbb{P}(Z_m = u) \to \frac{2\theta_u}{\sigma^2 p_0^u}$ as $m \to \infty$. Therefore,

$$\begin{aligned}
\lim_{n\to\infty} \frac{n^{2a} \mathbb{P}(Z_{n-k_1} = k_2)}{n^2 \mathbb{P}(Z_n = k_2)} &= \lim_{n\to\infty} \frac{\frac{2\theta_{k_2}}{\sigma^2 p_0^{k_2}}}{\frac{2\theta_{k_2}}{\sigma^2 p_0^{k_2}}} \\
&= 1
\end{aligned} \tag{4.48}$$

Finally, we turn our attention to the asymptotic behavior of the conditional expectation

$$\mathbb{E}\bigg(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} n^{2(1-a)} \mid Z_{k_1} > 0\bigg). \tag{4.49}$$

From (4.23),

$$
\begin{aligned}
q_{n-k_1}^{Z_{k_1}-1} &= \exp\left[(Z_{k_1}-1)(1-q_{n-k_1})\frac{\ln\left(1-(1-q_{n-k_1})\right)}{1-q_{n-k_1}}\right] \\
&= \exp\left[\frac{Z_{k_1}-1}{k_1}\cdot\frac{\ln\left(1-(1-q_{n-k_1})\right)}{1-q_{n-k_1}}\cdot(n-k_1)(1-q_{n-k_1})\cdot\frac{k_1}{n-k_1}\right].
\end{aligned} \tag{4.50}
$$

Each term may be considered separately in the following work, using the fact that $n-k_1 \to \infty$ and $k_1 \to \infty$ as $n \to \infty$.

First,

$$
\begin{aligned}
\lim_{n\to\infty}\frac{Z_{k_1}-1}{k_1} &= \lim_{n\to\infty}\frac{Z_{k_1}}{k_1}-\frac{1}{k_1} \\
&= \frac{\sigma^2}{2}\eta,
\end{aligned} \tag{4.51}
$$

in distribution under $\mathbb{P}_{\{Z_{k_1}>0\}}$ where η is an exponentially distributed random variable with parameter 1. Next, $\lim_{n\to\infty} n-k_1 = \lim_{n\to\infty} n^a = \infty$ for $a \in (0,1)$, implying that $\lim_{n\to\infty} 1-q_{n-k_1} = 0$.

Then,

$$
\lim_{n\to\infty}\frac{\ln\left(1-(1-q_{n-k_1})\right)}{1-q_{n-k_1}} = -1, \tag{4.52}
$$

Since $\lim_{u\to 0}\frac{\ln(1-u)}{u} = -1$.

Then,

$$
\lim_{n\to\infty}(n-k_1)(1-q_{n-k_1}) = \frac{2}{\sigma^2}, \tag{4.53}
$$

as in (4.7).

Finally, we observe that

$$
\begin{aligned}
\lim_{n\to\infty}\frac{k_1}{n-k_1} &= \lim_{n\to\infty}\frac{n-n^a}{n^a} \\
&= \lim_{n\to\infty} n^{1-a}-1 \\
&= \infty,
\end{aligned} \tag{4.54}
$$

since $a \in (0,1)$. In particular, we will see that the rate of this divergence is crucial to the remainder of the proof. Combining (4.51)-(4.54) and returning to the equivalence in (4.50),

$$\begin{aligned}\lim_{n\to\infty} q_{n-k_1}^{Z_{k_1}-1} &= \lim_{n\to\infty} \exp\left[\frac{Z_{k_1}-1}{k_1}\cdot\frac{\ln\left(1-(1-q_{n-k_1})\right)}{1-q_{n-k_1}}\cdot(n-k_1)(1-q_{n-k_1})\cdot\frac{k_1}{n-k_1}\right] \\ &= 0,\end{aligned} \tag{4.55}$$

where this convergence behaves in the same way as $e^{-c\cdot n^{1-a}}$, which is to say the convergence to 0 is of exponential order.

We may now return to the conditional expectation from (4.49) and consider its asymptotic behavior under $\mathbb{P}_{\{Z_{k_1}>0\}}$. We have seen that $\frac{Z_{k_1}}{k_1}$ converges to an exponentially distributed random variable in the distributional sense under $\mathbb{P}_{\{Z_{k_1}>0\}}$. It remains to note that while $q_{n-k_1}^{Z_{k_1}-1}\to 0$ exponentially,

$$\lim_{n\to\infty} n^{2(1-a)} = \infty, \tag{4.56}$$

polynomially. Thus,

$$\mathbb{E}\left(\frac{Z_{k_1}}{k_1} q_{n-k_1}^{Z_{k_1}-1} n^{2(1-a)} \mid Z_{k_1} > 0\right) \to 0 \tag{4.57}$$

as $n\to\infty$. Combining the results from (4.47), (4.48), and (4.57) verifies that

$$\lim_{n\to\infty} \mathbb{P}(Y_n > k_1 \mid Z_n = k_2) = 0.$$

□

As a result of Theorems 4.2.4 and 4.2.5, we want to make some observations to appreciate what these theorems imply in relation to the asymptotic behavior of the total coalescence time in the critical branching process. Considering the coalescence time of the order n but strictly away from n is not enough to gain the full picture of the desired limit in the critical case. With the lower bound on total coalescence time behaving like $c_1 n$ as in Theorem 4.2.4, coalescence is almost surly between generation $c_1 n$ and the present time n for any $c_1 < 1$.

Further, with the bound on total coalescence time behaving asymptotically like $n - n^a$ for $a \in (0,1)$ as in Theorem 4.2.5, coalescence is almost surely between the initial generation and generation $n - n^a$. This shows that there is very fine scaling required in order to further restrict the region of coalescence and maintain a non-trivial limit in this case. In particular, we would need to investigate a new bounding function $g(n)$ that will lie between $c\,n$ and $n - n^a$ for n large enough in order to gain a fuller picture of this asymptotic behavior.

CHAPTER

5

ASYMPTOTIC BEHAVIOR OF THE PAIRWISE COALESCENCE TIME

The goal of this final chapter is to construct results analogous to those found in Chapter 4, but now for the so-called pairwise coalescence time.

Definition 5.0.1. *The pairwise coalescence time of two randomly chosen individuals in any generation is the generation number of their most recent common ancestor. This will be a random variable and we denote by X_n the pairwise coalescence time for two randomly chosen individuals from generation n.*

In order to observe how the random variable X_n behaves asymptotically, we must first obtain several results regarding a point process that will turn up in the main pairwise coalescence convergence results, Theorems 5.2.1 and 5.2.2. Of these two culminating results in Section 5.2, the reader may observe that the first theorem is previously known and the second is an original theorem that broadens the scope of the situation in question. It can also be seen that Theorem 5.2.1 may be obtained by considering a special case of Theorem 5.2.2.

5.1 Convergence of a Particular Point Process

We have already seen that each $\{\frac{Z_n}{n} \mid Z_n > 0\}$ converges to an exponential random variable in the distributional sense. In the following result from [1], we are considering a random number Z_k of these points to see if they behave in the same way.

5.1.1 Case I

Theorem 5.1.1. *Let $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. Let the point process V_n be defined by*

$$V_n \doteq \left\{ \frac{Z_{n-k,i}^{(k)}}{n-k} : 1 \le i \le Z_k,\ Z_{n-k,i}^{(k)} > 0 \right\}.$$

Conditioned on the event $\{Z_n > 0\}$, as $n,\ k \to \infty$, $\frac{k}{n} \to u$, and with $u \in (0,1)$, V_n converges to the point process given by

$$V \doteq \{\eta_i : 1 \le i \le N_u\},$$

where $\{\eta_i\}_{i \ge 1}$ are i.i.d. exponential random variables with parameter $\frac{\sigma^2}{2}$ and N_u is independent of $\{\eta_i\}_{i\ge1}$ with distribution $\mathbb{P}(N_u = k) = (1-u)u^{k-1}$, $k \ge 1$ with the intensity 1.

Remark: It is intuitive that we will still get convergence to exponential random variables when considering random variables of this form (based on Theorem 3.2.1), but the interest rests on the collective convergence of all of them as a point process governed by the random variable N_u.

Proof. We may use the Laplace transform to prove an equivalent statement (See Section 2.5.1). Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a bounded and continuous function such that $f(0) = 0$. If it can be shown that for any $s > 0$,

$$\mathbb{E}\Big[\exp\Big(-s\sum_{i=1}^{Z_k} f\Big(\frac{Z_{n-k,i}^{(k)}}{n-k}\Big)\cdot \chi_{\{Z_{n-k,i}^{(k)} > 0\}}\Big) \,\Big|\, Z_n > 0\Big] \longrightarrow \mathbb{E}\Big[\exp\Big(-s\sum_{i=1}^{N_u} f(\eta_i)\Big)\Big],$$

we conclude the desired result, that conditioned on $\{Z_n > 0\}$, as $n,\ k \to \infty$, $\frac{k}{n} \to u$, where $u \in (0,1)$,

$$\left\{ \frac{Z_{n-k,i}^{(k)}}{n-k} : 1 \le i \le Z_k,\ Z_{n-k,i}^{(k)} > 0 \right\} \longrightarrow \{\eta_i : 1 \le i \le N_u\},$$

in distribution, as point processes. For $1 \le k \le n$, let

$$Y_{n,k} \doteq \exp\Big(-s\sum_{i=1}^{Z_k} f\Big(\frac{Z^{(k)}_{n-k,i}}{n-k}\Big)\cdot \chi_{\{Z^{(k)}_{n-k,i}\,>\,0\}}\Big).$$

We must look at $\mathbb{E}(Y_{n,k} \mid Z_n > 0)$. First we use the tower property for the conditional expectation to see that

$$\mathbb{E}(Y_{n,k} \mid Z_n > 0) = \mathbb{E}\big[\mathbb{E}(Y_{n,k}\cdot \chi_{\{Z_n\,>\,0\}} \mid \mathscr{F}_k) \mid Z_n > 0\big], \tag{5.1}$$

where $\mathscr{F}_k \doteq \sigma(\{Z_j \mid j \le k\})$, the σ-algebra generated by the set $\{Z_j \mid j \le k\}$. We will return to this equation, but first, consider the inside conditional expectation. We compute

$$\begin{aligned}
\mathbb{E}(Y_{n,k}\cdot\chi_{\{Z_n>0\}} \mid \mathscr{F}_k) &= \mathbb{E}\big(Y_{n,k}\cdot\chi_{\{Z_k>0\}}\cdot\chi_{\{Z_n>0\}} \mid Z_k\big) \qquad \text{(by the Markov Property)}\\
&= \mathbb{E}\big(Y_{n,k}\cdot\chi_{\{Z_k>0\}} \mid Z_k\big) - \mathbb{E}\big(Y_{n,k}\cdot\chi_{\{Z_k>0\}}\cdot\chi_{\{Z_n=0\}} \mid Z_k\big)\\
&= \mathbb{E}\Big[\exp\Big(-s\sum_{i=1}^{Z_k} f\Big(\frac{Z^{(k)}_{n-k,i}}{n-k}\Big)\cdot\chi_{\{Z^{(k)}_{n-k,i}>0\}}\Big)\cdot\chi_{\{Z_k>0\}} \mid Z_k\Big]\\
&\qquad -\mathbb{E}\Big[\exp\Big(-s\sum_{i=1}^{Z_k} f\Big(\frac{Z^{(k)}_{n-k,i}}{n-k}\Big)\cdot\chi_{\{Z^{(k)}_{n-k,i}>0\}}\Big)\cdot\chi_{\{Z_k>0\}}\cdot\chi_{\{Z_n=0\}} \mid Z_k\Big]\\
&= \Big[\mathbb{E}\Big(\exp\big(-sf(\frac{Z_{n-k}}{n-k})\cdot\chi_{\{Z_{n-k}>0\}}\big) \mid Z_0=1\Big)\Big]^{Z_k}\cdot\chi_{\{Z_k>0\}}\\
&\qquad -\Big[\mathbb{P}\big(Z_{n-k}=0 \mid Z_0=1\big)\Big]^{Z_k}\cdot\chi_{\{Z_k>0\}} \qquad \text{(Due to conditional independence)}\\
&= \big(g_{n-k}(s)\big)^{Z_k}\cdot\chi_{\{Z_k>0\}} - \big(q_{n-k}\big)^{Z_k}\cdot\chi_{\{Z_k>0\}},
\end{aligned}$$

where $g_j(s) \doteq \mathbb{E}\big[\exp\big(-sf(\frac{Z_j}{j})\cdot\chi_{\{Z_j\,>\,0\}}\big) \mid Z_0 = 1\big]$ and $q_j \doteq \mathbb{P}\big(Z_j = 0 \mid Z_0 = 1\big)$.
In summary, from above,

$$\mathbb{E}(Y_{n,k}\cdot\chi_{\{Z_n>0\}} \mid \mathscr{F}_k) = \big(g_{n-k}(s)\big)^{Z_k}\cdot\chi_{\{Z_k>0\}} - \big(q_{n-k}\big)^{Z_k}\cdot\chi_{\{Z_k>0\}}. \tag{5.2}$$

Let $\tilde{g}_j(s) \doteq \mathbb{E}\big[\exp\big(-sf(\frac{Z_j}{j})\big) \mid Z_j > 0,\, Z_0 = 1\big]$. Recall from Theorem 3.2.1 that $\big\{\frac{Z_j}{j} \mid Z_j > 0\big\} \xrightarrow{d} Y$ where Y is exponentially distributed with parameter $\frac{2}{\sigma^2}$. Then since f is a bounded and continuous function, we observe the following as $j \to \infty$.

$$\begin{aligned}\tilde{g}_j(s) &= \int_{\mathbb{R}} e^{-sf(x)} d\big(\mathbb{P}_{\{Z_j>0\}}\big)_{\frac{Z_j}{j}}(x) \\ &\to \frac{2}{\sigma^2}\int_0^\infty e^{-sf(x)} e^{\frac{-2x}{\sigma^2}}\, dx \\ &\doteq \tilde{g}(s).\end{aligned} \tag{5.3}$$

Now, if $g_j(s)$ can be expressed in terms of $\tilde{g}_j(s)$, then we may observe its limiting behavior. Note that

$$\begin{aligned}\tilde{g}_j(s) &= \mathbb{E}\Big[\exp\big(-sf(\frac{Z_j}{j})\big) \mid Z_j > 0,\ Z_0 = 1\Big] \\ &= \sum_{k=1}^{\infty} \exp\big(-sf(k/j)\big)\mathbb{P}(Z_j = k \mid Z_j > 0,\ Z_0 = 1) \\ &= \sum_{k=1}^{\infty} \exp\big(-sf(k/j)\big)\mathbb{P}(Z_j = k,\ Z_0 = 1)\cdot\frac{1}{\mathbb{P}(Z_j > 0,\ Z_0 = 1)} \\ &= \sum_{k=1}^{\infty} \exp\big(-sf(k/j)\big)\mathbb{P}(Z_j = k,\ Z_0 = 1)\cdot\frac{1}{P(Z_j > 0 \mid Z_0 = 1)\cdot\mathbb{P}(Z_0 = 1)} \\ &= \sum_{k=1}^{\infty} \exp\big(-sf(k/j)\big)\mathbb{P}(Z_j = k \mid Z_0 = 1)\cdot\frac{1}{(1-\mathbb{P}(Z_j = 0 \mid Z_0 = 1))} \\ &= \frac{g_j(s) - e^{-sf(0)}q_j}{1-q_j} \\ &= \frac{g_j(s) - q_j}{1-q_j},\end{aligned}$$

where we used the fact that $f(0) = 0$. Then,

$$\begin{aligned}g_j(s) &= q_j + \tilde{g}_j(s)(1-q_j) \\ &= 1 + (1-q_j)(\tilde{g}_j(s)-1).\end{aligned} \tag{5.4}$$

Now that $g_j(s)$ is known in terms of $\tilde{g}_j(s)$, we are able to determine the limiting behavior of not just $g_j(s)$, but $(g_j(s))^j$, as well as q_j^j. First, see the following equivalences:

$$\begin{aligned}(g_j(s))^j &= \big(1 + (1-q_j)(\tilde{g}_j(s)-1)\big)^j \\ &= \exp\big[j\ln(1 + (1-q_j)(\tilde{g}_j(s)-1))\big] \\ &= \exp\Big[j(1-q_j)\frac{\ln(1 + (1-q_j)(\tilde{g}_j(s)-1))}{1-q_j}\Big],\end{aligned} \tag{5.5}$$

and

$$
\begin{aligned}
q_j^j &= \exp\left[j \ln(1-(1-q_j))\right] \\
&= \exp\left[j(1-q_j)\frac{\ln(1-(1-q_j))}{1-q_j}\right].
\end{aligned} \tag{5.6}
$$

Computing the limits of (5.5) and (5.6) respectively as $j \to \infty$ yields

$$
\begin{aligned}
\lim_{j\to\infty} (g_j(s))^j &= \exp\left[2/\sigma^2 \lim_{j\to\infty} \frac{\ln(1 + (1-q_j)(\tilde{g}_j(s)-1))}{1-q_j}\right] \\
&= \exp\left[2/\sigma^2 \lim_{j\to\infty} (\tilde{g}_j(s)-1)\right] \\
&= \exp\left[\frac{2}{\sigma^2}(\tilde{g}(s)-1)\right],
\end{aligned} \tag{5.7}
$$

recalling from (5.3) that $\tilde{g}_j(s) \to \tilde{g}(s)$ as $j \to \infty$. Also,

$$
\begin{aligned}
\lim_{j\to\infty} q_j^j &= \exp\left[2/\sigma^2 \lim_{j\to\infty} \frac{\ln(1-(1-q_j))}{1-q_j}\right] \\
&= \exp\left[-2/\sigma^2\right].
\end{aligned} \tag{5.8}
$$

Returning to (5.1) and using (5.2),

$$
\mathbb{E}(Y_{n,k} \mid Z_n > 0) = \mathbb{E}\left[\left(\left(g_{n-k}(s)\right)^{Z_k} \cdot \chi_{\{Z_k > 0\}} - \left(q_{n-k}\right)^{Z_k} \cdot \chi_{\{Z_k > 0\}}\right) \mid Z_n > 0\right]. \tag{5.9}
$$

Consider the following to determine the limiting behavior of the two parts of the right-hand term in (5.9) as $n \to \infty$, $k \to \infty$, $\frac{k}{n} \to u$, and with $u \in (0,1)$.

$$\begin{aligned}
\mathbb{E}\Big[\big(g_{n-k}(s)\big)^{Z_k} \;\Big|\; Z_k > 0\Big] &= \mathbb{E}\Big[\Big(\big(g_{n-k}(s)\big)^{n-k}\Big)^{\frac{Z_k}{k}(\frac{k}{n-k})} \;\Big|\; Z_k > 0\Big] \\
&\to \frac{2}{\sigma^2}\int_0^\infty \exp\Big[\frac{2}{\sigma^2}(\tilde{g}(s)-1)\cdot\frac{u}{1-u}\cdot x\Big]e^{\frac{-2x}{\sigma^2}}\,dx,
\end{aligned}$$

recalling the limit computed in (5.5) and that from Theorem 3.2.1, $\frac{Z_k}{k}$ under $\mathbb{P}_{\{Z_k>0\}}$ converges in distribution to Y where Y is an exponentially distributed random variable with parameter $\frac{2}{\sigma^2}$.

$$\begin{aligned}
&= \frac{2}{\sigma^2}\int_0^\infty \exp\Big[\frac{-2}{\sigma^2}\big(1-\frac{(\tilde{g}(s)-1)u}{1-u}\big)x\Big]\,dx \\
&= \frac{2}{\sigma^2}\cdot\frac{\sigma^2}{2\big(1-\frac{(\tilde{g}(s)-1)u}{1-u}\big)} \\
&= \frac{1-u}{1-u\tilde{g}(s)}.
\end{aligned} \tag{5.10}$$

Also,

$$\begin{aligned}
\mathbb{E}\Big[(q_{n-k})^{Z_k} \;\Big|\; Z_k > 0\Big] &= \mathbb{E}\Big[\big((q_{n-k})^{n-k}\big)^{\frac{Z_k}{k}\cdot\frac{k}{n-k}} \;\Big|\; Z_k > 0\Big] \\
&\to \frac{2}{\sigma^2}\int_0^\infty \exp\Big[\frac{-2}{\sigma^2}\cdot\frac{u}{1-u}\cdot x\Big]e^{\frac{-2x}{\sigma^2}}\,dx,
\end{aligned}$$

recalling the limit computed in (5.6).

$$\begin{aligned}
&= \frac{2}{\sigma^2}\int_0^\infty \exp\Big[\frac{-2}{\sigma^2}\Big(\frac{u}{1-u}+1\Big)x\Big]\,dx \\
&= 1-u.
\end{aligned} \tag{5.11}$$

From the (5.9), however, we must consider these expectations conditioned not on $\{Z_k > 0\}$ but rather on $\{Z_n > 0\}$. To that end, we use (5.10) and (5.11) to compute the following, applying Theorem 3.1.1 whenever necessary. As $n \to \infty$, $k \to \infty$, $\frac{k}{n} \to u$, and with $u \in (0,1)$,

$$\begin{aligned}\mathbb{E}\Big[\big(g_{n-k}(s)\big)^{Z_k} \mid Z_n > 0\Big] &= \frac{\mathbb{P}(Z_k > 0)}{\mathbb{P}(Z_n > 0)} \cdot \mathbb{E}\Big[\big(g_{n-k}(s)\big)^{Z_k} \;\Big|\; Z_k > 0\Big] \\ &\to \frac{1}{u} \cdot \frac{1-u}{1-u\tilde{g}(s)},\end{aligned} \tag{5.12}$$

recalling from Theorem 3.1.1 that $m\mathbb{P}(Z_m > 0) \to \frac{2}{\sigma^2}$. Similarly, and under the same limiting considerations,

$$\begin{aligned}\mathbb{E}\Big[(q_{n-k})^{Z_k} \mid Z_n > 0\Big] &= \lim_{n\to\infty} \frac{\mathbb{P}(Z_k > 0)}{\mathbb{P}(Z_n > 0)} \cdot \mathbb{E}\Big[(q_{n-k})^{Z_k} \;\Big|\; Z_k > 0\Big] \\ &\to \frac{1}{u}(1-u).\end{aligned} \tag{5.13}$$

Finally, we utilize (5.12) and (5.13) by returning to (5.9) and considering the asymptotic behavior thereof as $n \to \infty$, $k \to \infty$, $\frac{k}{n} \to u$, and with $u \in (0,1)$.

$$\begin{aligned}\mathbb{E}(Y_{n,k} \mid Z_n > 0) &\to \frac{1-u}{u}\Big[\frac{1}{1-u\tilde{g}(s)} - 1\Big] \\ &= (1-u) \cdot \frac{\tilde{g}(s)}{1-u\tilde{g}(s)} \\ &= (1-u)\sum_{j=0}^{\infty} u^j\big(\tilde{g}(s)\big)^{j+1} \\ &= \sum_{j=1}^{\infty}(1-u)u^{j-1}\big(\tilde{g}(s)\big)^j \\ &= \sum_{j=1}^{\infty}\mathbb{P}(N_u = j)\Big[\frac{2}{\sigma^2}\int_0^{\infty} e^{-sf(x)}e^{\frac{-2x}{\sigma^2}}\,dx\Big]^j \\ &= \sum_{j=1}^{\infty}\mathbb{P}(N_u = j)\mathbb{E}\Big(e^{-s\sum_{i=1}^{j} f(\eta_i)}\Big).\end{aligned} \tag{5.14}$$

Here we require the linear independence between the random variable N_u and the family of exponential random variables $\{\eta_i\}_{i\geq 1}$ to see that

$$\sum_{j=1}^{\infty}\mathbb{P}(N_u = j)\mathbb{E}\Big(e^{-s\sum_{i=1}^{j} f(\eta_i)}\Big) = \mathbb{E}\Big(e^{-s\sum_{i=1}^{N_u} f(\eta_i)}\Big). \tag{5.15}$$

Then, combining (5.14) and (5.15), as $n \to \infty$, $k \to \infty$, $\frac{k}{n} \to u$, and with $u \in (0,1)$,

$$\mathbb{E}(Y_{n,k} \mid Z_n > 0) \to \mathbb{E}\Big(e^{-s\sum_{i=1}^{N_u} f(\eta_i)}\Big),$$

where the exponential random variables $\{\eta_i\}_{i\geq 1}$ are i.i.d. with parameter $\frac{\sigma^2}{2}$ and N_u is geometrically distributed with parameter u and independent of $\{\eta_i\}_{i\geq 1}$. $\square$

5.1.2 Case II

We now present an adjustment to the previous theorem as a result of our own consideration. This non-trivial change to the conditions considered yields a new distribution function with intensity dependent on a given constant which arises out of the limit in question.

Theorem 5.1.2. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Let the point process* V_n *be defined by*

$$V_n \doteq \left\{ \frac{Z^{(k)}_{n-k_1,i}}{n-k_1} \,:\, 1 \leq i \leq Z_{k_1},\, Z^{(k_1)}_{n-k_1,i} > 0 \right\}.$$

Then conditioned on the event $\{Z_n > k_2\}$, *as* $n,\ k_1,\ k_2 \to \infty$, $\frac{k_1}{n} \to u_1$, $\frac{k_2}{n} \to u_2$, *and with* $u_1 \in (0,1)$ *and* $u_2 > 0$, V_n *converges to the point process*

$$V \doteq \{\eta_i \,:\, 1 \leq i \leq N_{u_1}\},$$

where $\{\eta_i\}_{i\geq 1}$ *are i.i.d. exponential random variables with parameter* $\frac{\sigma^2}{2}$ *and* N_{u_1} *is independent of* $\{\eta_i\}_{i\geq 1}$ *with distribution* $\mathbb{P}(N_{u_1} = k) = (1 - u_1)u_1^{k-1}$, $k \geq 1$, *with intensity given by* $e^{\frac{2u_2}{\sigma^2}}$.

Proof. We may again use the Laplace transform to prove an equivalent statement. Let $f : \mathbb{R}^+ \to \mathbb{R}^+$ be a bounded and continuous function such that $f(0) = 0$. If it can be shown that for any $s > 0$,

$$\mathbb{E}\left[\exp\left(-s\sum_{i=1}^{Z_{k_1}} f\left(\frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)\cdot \chi_{\{Z^{(k_1)}_{n-k_1,i} > 0\}}\right) \,\Big|\, Z_n > k_2\right] \longrightarrow \mathbb{E}\left[\exp\left(-s\sum_{i=1}^{N_{u_1}} f(\eta_i)\right)\right],$$

we conclude the desired result, that conditioned on $\{Z_n > k_2\}$, as $n,\ k \to \infty$, $\frac{k_1}{n} \to u_1$, $\frac{k_2}{n} \to u_2$, where $u_1 \in (0,1)$, $u_2 > 0$,

$$\left\{ \frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1} \,:\, 1 \leq i \leq Z_{k_1},\, Z^{(k_1)}_{n-k_1,i} > 0 \right\} \longrightarrow \{\eta_i \,:\, 1 \leq i \leq N_{u_1}\},$$

in distribution, as point processes. For $1 \leq k_1 \leq n$, let

$$Y_{n,k_1} \doteq \exp\left(-s\sum_{i=1}^{Z_{k_1}} f\left(\frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)\cdot \chi_{\{Z^{(k_1)}_{n-k_1,i} > 0\}}\right).$$

We must look at $\mathbb{E}(Y_{n,k_1} \mid Z_n > k_2)$. As in the proof of Theorem 5.1.1, we use the tower property to see that

$$\mathbb{E}(Y_{n,k_1} \mid Z_n > k_2) = \mathbb{E}\big[\mathbb{E}(Y_{n,k_1} \cdot \chi_{\{Z_n > 0\}} \mid \mathscr{F}_{k_1}) \mid Z_n > k_2\big], \tag{5.16}$$

where $\mathscr{F}_{k_1} \doteq \sigma(\{Z_j \mid j \leq k_1\})$, the σ-algebra generated by the set $\{Z_j \mid j \leq k_1\}$. Recall (5.2), which says that

$$\mathbb{E}(Y_{n,k_1} \cdot \chi_{\{Z_n>0\}} \mid \mathscr{F}_{k_1}) = \big(g_{n-k_1}(s)\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} - \big(q_{n-k_1}\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}}, \tag{5.17}$$

where $g_j(s) \doteq \mathbb{E}\Big[\exp\big(-sf(\frac{Z_j}{j}) \cdot \chi_{\{Z_j > 0\}}\big) \mid Z_0 = 1\Big]$ and $q_j \doteq \mathbb{P}\big(Z_j = 0 \mid Z_0 = 1\big)$. Note that this equivalence still holds because in Theorem 5.1.1, $\frac{k}{n} \to u$ as $n \to \infty$ whereas here, $\frac{k_1}{n} \to u_1$ as $n \to \infty$.

Thus, (5.16) and (5.17) together yield

$$\begin{aligned}\mathbb{E}(Y_{n,k_1} \mid Z_n > k_2) &= \mathbb{E}\Big[\big(g_{n-k_1}(s)\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} - \big(q_{n-k_1}\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} \,\Big|\, Z_n > k_2\Big] \\ &= \mathbb{E}\Big[\big(g_{n-k_1}(s)\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} \,\Big|\, Z_n > k_2\Big] - \mathbb{E}\Big[\big(q_{n-k_1}\big)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} \,\Big|\, Z_n > k_2\Big].\end{aligned} \tag{5.18}$$

As in Theorem 5.1.1, let $\tilde{g}_j(s) \doteq \mathbb{E}\Big[\exp\big(-sf(\frac{Z_j}{j})\big) \mid Z_j > 0,\ Z_0 = 1\Big]$ and note that (5.3) computed $\lim_{j\to\infty} \tilde{g}_j(s) \doteq \tilde{g}(s)$. (5.10) and (5.11) respectively show that as $n,\ k_1 \to \infty$, $\frac{k_1}{n} \to u_1$, and with $u_1 \in (0,1)$,

$$\mathbb{E}\Big[\big(g_{n-k_1}(s)\big)^{Z_{k_1}} \,\Big|\, Z_{k_1} > 0\Big] \to \frac{1 - u_1}{1 - u_1\tilde{g}(s)}, \tag{5.19}$$

and

$$\mathbb{E}\Big[(q_{n-k_1})^{Z_{k_1}} \,\Big|\, Z_{k_1} > 0\Big] \to 1 - u_1. \tag{5.20}$$

From the (5.18), however, we must consider these expectations conditioned not on $\{Z_{k_1} > 0\}$ but rather on $\{Z_n > k_2\}$. To that end, we use (5.19) and (5.20) to compute the following, applying Theorem 3.1.1 whenever necessary and also using Theorem 3.2.1 to compute $\lim_{n\to\infty} \mathbb{P}(Z_n > k_2)$ in the present case. As $n,\ k_1,\ k_2 \to \infty$, $\frac{k_1}{n} \to u_1$, $\frac{k_2}{n} \to u_2$, and with

$u_1 \in (0,1),\ u_2 > 0,$

$$\begin{aligned}\mathbb{E}\left[\left(g_{n-k_1}(s)\right)^{Z_{k_1}} \cdot \chi_{\{Z_{k_1}>0\}} \,\middle|\, Z_n > k_2\right] &= \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \cdot \mathbb{E}\left[\left(g_{n-k_1}(s)\right)^{Z_{k_1}} \,\middle|\, Z_{k_1} > 0\right] \\ &\to \frac{1}{u_1} e^{\frac{2u_2}{\sigma^2}} \cdot \frac{1-u_1}{1-u_1\tilde{g}(s)}.\end{aligned} \tag{5.21}$$

Similarly,

$$\begin{aligned}\mathbb{E}\left[(q_{n-k_1})^{Z_{k_1}} \mid Z_n > k_2\right] &= \frac{\mathbb{P}(Z_{k_1} > 0)}{\mathbb{P}(Z_n > k_2)} \cdot \mathbb{E}\left[(q_{n-k_1})^{Z_{k_1}} \,\middle|\, Z_{k_1} > 0\right] \\ &\to \frac{1}{u_1} e^{\frac{2u_2}{\sigma^2}} (1-u_1).\end{aligned} \tag{5.22}$$

Finally, we utilize (5.21) and (5.22) by returning to (5.18) and considering the asymptotic behavior thereof as $n,\ k_2,\ k_2 \to \infty,\ \frac{k_1}{n} \to u_1,\ \frac{k_2}{n} \to u_2$ and with $u_1 \in (0,1),\ u_2 > 0.$

$$\begin{aligned}\mathbb{E}(Y_{n,k_1} \mid Z_n > k_2) &\to \frac{1}{u_1} e^{\frac{2u_2}{\sigma^2}} \left[\frac{1-u_1}{1-u_1\tilde{g}(s)} - (1-u_1)\right] \\ &= (1-u_1) e^{\frac{2u_2}{\sigma^2}} \cdot \frac{\tilde{g}(s)}{1-u_1\tilde{g}(s)} \\ &= (1-u_1) e^{\frac{2u_2}{\sigma^2}} \sum_{j=0}^{\infty} u_1^j \left(\tilde{g}(s)\right)^{j+1} \\ &= \sum_{j=1}^{\infty} (1-u_1) u_1^{j-1} e^{\frac{2u_2}{\sigma^2}} \left(\tilde{g}(s)\right)^j \\ &= \sum_{j=1}^{\infty} e^{\frac{2}{\sigma^2}u_2} \mathbb{P}(N_{u_1} = j) \left[\frac{2}{\sigma^2} \int_0^{\infty} e^{-sf(x)} e^{\frac{-2x}{\sigma^2}}\, dx\right]^j \\ &= \sum_{j=1}^{\infty} e^{\frac{2}{\sigma^2}u_2} \mathbb{P}(N_{u_1} = j) \mathbb{E}\left[e^{-s\sum_{i=1}^{j} f(\eta_i)}\right].\end{aligned} \tag{5.23}$$

This is the characteristic function for the point process V defined above by

$$V = \{\eta_i : 1 \le i \le N_{u_1}\},$$

where $\{\eta_i\}_{i\ge1}$ are i.i.d. exponential random variables with parameter $\frac{\sigma^2}{2}$ and N_{u_1} is independent of $\{\eta_i\}_{i\ge1}$ with distribution $\mathbb{P}(N_{u_1} = k) = (1-u_1)u_1^{k-1},\ k \ge 1$, with intensity given by $e^{\frac{2u_2}{\sigma^2}}$. □

5.2 Main Pairwise Coalescence Results

The results given in Section 5.1 are particularly useful in that they contribute to gaining a fuller understanding of the asymptotic behavior of the pairwise coalescence time under multi-scale conditioning. Recall the definition of pairwise coalescence given in Definition (5.0.1). Theorem (5.2.1) comes from Athreya's work in [1] and gives way to an original result based on adjusted conditioning in Theorem 5.2.2 that follows.

5.2.1 Case I

Theorem 5.2.1. *Let* $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, *and* $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. *Then for* $0 < u < 1$,

$$\lim_{n\to\infty} \mathbb{P}(X_n < nu \mid Z_n > 1) \equiv H(u)$$

where $H(u)$ *exists and*

$$H(u) = 1 - \mathbb{E}(\phi(N_u)) \tag{5.24}$$

such that:

- N_u *is a positive, integer-valued, geometrically-distributed random variable, i.e.* $\mathbb{P}(N_u = k) = (1-u)u^{k-1}$ *for all* $k \geq 1$;
- *for* $j \geq 1$,

$$\phi(j) = \mathbb{E}\left(\frac{\sum_{i=1}^{j} \eta_i^2}{\left(\sum_{i=1}^{j} \eta_i\right)^2}\right); \tag{5.25}$$

- $\{\eta_i\}_{i\geq 0}$ *are i.i.d. exponential random variables with mean* 1.

Furthermore, $H(\cdot)$ *is an absolutely continuous cumulative distribution function on* $[0,1]$ *with* $\lim_{u\to 0^+} H(u) = 0$ *and* $\lim_{u\to 1^-} H(u) = 1$.

Proof. First, note that the event $\{X_n < k\}$ occurs if and only if the two randomly chosen individuals from generation n are offspring of 2 distinct individuals of generation k. Then these two individuals are part of the $(n-k)$-th generation of trees starting by two distinct individuals in generation k. This is due to analogous reasoning to that of considering the event $\{Y_n > k\}$ in Theorems 4.1.1-4.1.5. This yields

$$\begin{aligned}\mathbb{P}(X_n < k \mid Z_n > 1) &= \mathbb{E}\Bigg[\frac{\sum\limits_{1\le i\ne j\le Z_k} Z^{(k)}_{n-k,i}Z^{(k)}_{n-k,j}}{Z_n(Z_n-1)} \;\Big|\; Z_n > 1\Bigg] \\ &= \mathbb{E}\Bigg[\frac{\big(\sum_{i=1}^{Z_k} Z^{(k)}_{n-k,i}\big)^2 - \sum_{i=1}^{Z_k}(Z^{(k)}_{n-k,i})^2}{Z_n(Z_n-1)} \;\Big|\; Z_n > 1\Bigg].\end{aligned} \tag{5.26}$$

Note that each of the individuals in generation n must come from a branching process initiated by one of the Z_k individuals in generation k. Thus,

$$\sum_{i=1}^{Z_k} Z^{(k)}_{n-k,i} = Z_n. \tag{5.27}$$

We also observe the following as $n \to \infty$:

$$\begin{aligned}\mathbb{E}(Z_n \mid Z_n > 0) \text{ grows in the order } n \;&\Rightarrow\; \left\{\frac{1}{Z_n} \;\Big|\; Z_n > 1\right\} \to 0 \\ &\Rightarrow\; \left\{1 - \frac{1}{Z_n} \;\Big|\; Z_n > 1\right\} \to 1 \\ &\Rightarrow\; \left\{\frac{Z_n(Z_n-1)}{Z_n^2} \;\Big|\; Z_n > 1\right\} \to 1.\end{aligned}$$

This confirms that under the present conditioning, $\frac{1}{Z_n(Z_n-1)}$ may be safely replaced by $\frac{1}{Z_n^2}$. In light of (5.27) and the above observation, (5.26) may be rewritten as

$$\begin{aligned}\mathbb{P}(X_n < k \mid Z_n > 1) &= \mathbb{E}\Bigg[\frac{Z_n^2 - \sum_{i=1}^{Z_k}(Z^{(k)}_{n-k,i})^2}{Z_n^2} \;\Big|\; Z_n > 1\Bigg] \\ &= \mathbb{E}\Bigg[1 - \frac{\sum_{i=1}^{Z_k}(Z^{(k)}_{n-k,i})^2}{\big(\sum_{i=1}^{Z_k} Z^{(k)}_{n-k,i}\big)^2} \;\Big|\; Z_n > 1\Bigg].\end{aligned} \tag{5.28}$$

We want to show that the latter term in (5.26) converges to $H(u)$ as $n \to \infty$ and concurrently $\frac{k}{n} \to u$. In order to see this, (5.28) shows that we may equivalently prove

$$\mathbb{E}\Bigg[\frac{\sum_{i=1}^{Z_k}(Z^{(k)}_{n-k,i})^2}{\big(\sum_{i=1}^{Z_k} Z^{(k)}_{n-k,i}\big)^2} \;\Big|\; Z_n > 1\Bigg] \longrightarrow 1 - H(u) = \mathbb{E}(\phi(N_u)), \tag{5.29}$$

as $n \to \infty$ and $\frac{k}{n} \to u$. Note that there may be some individuals from generation k who have "died out" before getting to generation n. To this end, let G_k be the set of all individuals alive in the generation k, then let $J_k \doteq \{i \in G_k \;:\; Z^{(k)}_{n-k,i} > 0\}$, the set of individuals in the generation k still alive at generation n. Then

$$\frac{\sum_{i=1}^{Z_k}(Z_{n-k,i}^{(k)})^2}{\left(\sum_{i=1}^{Z_k} Z_{n-k,i}^{(k)}\right)^2} = \frac{\sum_{i\in J_k}(\frac{Z_{n-k,i}^{(k)}}{n-k})^2}{\left(\sum_{i\in J_k}\frac{Z_{n-k,i}^{(k)}}{n-k}\right)^2}. \tag{5.30}$$

In order to consider the asymptotic behavior of the point processes given here, we seek to utilize Theorem 5.1.1. J_k is a continuous functional of the point process $\{V_n\}$ as defined in Theorem 5.1.1 and hence weak convergence obtained in the theorem would suffice to give limiting behavior in this case. One key point, however, is that Theorem 5.1.1 considers the point process conditioned on the event $\{Z_n > 0\}$ and in (5.30), we must consider convergence of the point process under the event $\{Z_n > 1\}$. To this end we compute the following, bearing in mind that from Theorem 3.2.1, $\frac{Z_n}{n}$ under $\mathbb{P}_{\{Z_n>0\}}$ converges in distribution to Y where Y is an exponentially distributed random variable with parameter $\frac{2}{\sigma^2}$.

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}(Z_n > 1 \mid Z_n > 0) &= \lim_{n\to\infty}\mathbb{P}\Big(\frac{Z_n}{n} > \frac{1}{n} \mid Z_n > 0\Big)\\ &= \lim_{n\to\infty}\mathbb{P}\big(Y > \frac{1}{n}\big) \text{ where Y has distribution } e^{-2u/\sigma^2} \text{ by Theorem 3.2.1.}\\ &= \mathbb{P}(Y > 0)\\ &= 1.\end{aligned} \tag{5.31}$$

As a result, we may utilize the Theorem 5.1.1 to obtain the desired convergence, namely,

$$\frac{\sum_{i\in J_k}(\frac{Z_{n-k,i}^{(k)}}{n-k})^2}{\left(\sum_{i\in J_k}\frac{Z_{n-k,i}^{(k)}}{n-k}\right)^2} \longrightarrow \frac{\sum_{i=1}^{N_u}\eta_i^2}{\left(\sum_{i=1}^{N_u}\eta_i\right)^2}, \tag{5.32}$$

in the distributional sense as $n\to\infty$ and $\frac{k}{n}\to u$ where $\{\eta_i\}_{i\geq 0}$ and N_u are as given in the statement of Theorem 5.1.1.

It remains only to see that $H(u)$ satisfies the desired limiting behavior. We have

$$H(u) = 1-\mathbb{E}(\phi(N_u)).$$

As $u\to 1^-$, we see that $N_u\to\infty$ almost surely and thus we consider the following. As a result of applying the bounded convergence theorem to the sequence of absolutely bounded random variables $f_j \doteq \frac{\sum_{i=1}^{j}\eta_i^2}{\left(\sum_{i=1}^{j}\eta_i\right)^2}$,

$$\lim_{u\to 1^-}\phi(N_u) = \lim_{u\to 1^-}\mathbb{E}\left(\frac{\sum_{i=1}^{N_u}\eta_i^2}{\left(\sum_{i=1}^{N_u}\eta_i\right)^2}\right) = \mathbb{E}\left(\lim_{u\to 1^-}\frac{\sum_{i=1}^{N_u}\eta_i^2}{\left(\sum_{i=1}^{N_u}\eta_i\right)^2}\right). \tag{5.33}$$

Now, to observe the asymptotic behavior of f_j, we utilize the strong law of large numbers. Note that $\mathbb{E}(\eta_1) = \frac{\sigma^2}{2} > 0$ and $\mathbb{E}(\eta_i^2) < \infty$ and so we may apply the strong law of large numbers to obtain

(i) $\lim_{u\to 1^-}\frac{1}{N_u}\sum_{i=1}^{N_u}\eta_i = \frac{\sigma^2}{2}$;

(ii) $\lim_{u\to 1^-}\frac{1}{N_u}\sum_{i=1}^{N_u}\eta_i^2 = \left(\frac{\sigma^2}{2}\right)^2$.

Then

$$\lim_{j\to\infty} f_j = 0, \tag{5.34}$$

with probability 1, for which we needed to recall the independence of the random variables $\{\eta_i\}_{i\geq 1}$ and the fact that N_u is independent of $\{\eta_i\}_{i\geq 1}$. Together with (5.33), (5.34) concludes that

$$\lim_{u\to 1^-}\phi(N_u) = 0. \tag{5.35}$$

Thus,

$$\begin{aligned}\lim_{u\to 1^-}H(u) &= \lim_{u\to 1^-}(1-\mathbb{E}(\phi(N_u)))\\ &= 1-\lim_{u\to 1^-}\mathbb{E}(\phi(N_u))\\ &= 1.\end{aligned}$$

Finally, since

$$\begin{aligned}H(u) &= 1-\mathbb{E}(\phi(N_u))\\ &= 1-\sum_{j=1}^{\infty}\phi(j)\mathbb{P}(N_u = j)\\ &= 1-\sum_{j=1}^{\infty}\phi(j)(1-u)u^{j-1},\end{aligned} \tag{5.36}$$

we see that

$$\begin{aligned}\lim_{u\to 0^+} H(u) &= 1-\phi(1)\\ &= 1-\mathbb{E}\left(\frac{\eta_1^2}{\eta_1^2}\right)\\ &= 0.\end{aligned}$$

□

We observe that Theorem 5.2.1 equivalently states that $\{\frac{X_n}{n} \mid Z_n > 1\}$ converges in distribution to Y where Y has distribution function $H(u)$. Also note that the double expectation computation carried out in $H(u)$ is on account of the randomness of both the sequence of random variables $\{\eta_i\}_{i\geq 1}$ and that of N_u. The expectation in (5.25) kills the randomness of the random variables $\{\eta_i\}_{i\geq 1}$, while that of N_u is handled by the expectation in (5.24).

5.2.2 Case II

Theorem 5.2.2. *Let $m \doteq \sum_{j=1}^{\infty} j p_j = 1$, $p_1 < 1$, and $\sigma^2 \doteq \sum_{j=1}^{\infty} j^2 p_j - 1 < \infty$. Then for $u_1 \in (0,1)$ and $u_2 > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(X_n < n u_1 \mid Z_n > n u_2) \equiv G(u_1, u_2)$$

where $G(u_1,u_2)$ exists and

$$G(u_1, u_2) = 1-\mathbb{E}(\phi(V)) \tag{5.37}$$

such that:

- *V is the point process generated by $\{\eta_i\}_{i\geq 1}$ and N_{u_1} with intensity given by $e^{\frac{2u_2}{\sigma^2}}$;*
- *$\{\eta_i\}_{i\geq 1}$ are i.i.d. exponential random variables with parameter $\frac{\sigma^2}{2}$ and N_{u_1} is independent of $\{\eta_i\}_{i\geq 1}$ with distribution $\mathbb{P}(N_{u_1} = k) = (1-u_1)u_1^{k-1}$, $k \geq 1$;*
- *for $j \geq 1$,*

$$\phi(j) = \mathbb{E}\left(\frac{\sum_{i=1}^{j} \eta_i^2}{\left(\sum_{i=1}^{j} \eta_i\right)^2}\right). \tag{5.38}$$

Proof. As in the proof of Theorem 5.2.1, note that the event $\{X_n < k\}$ occurs if and only if the two randomly chosen individuals from generation n are offspring of 2 distinct individuals of generation k. Then these two individuals are part of the $(n-k)$-th generation of trees starting by two distinct individuals in generation k. This is due to analogous reasoning to that of considering the event $\{Y_n > k\}$ in Theorems 4.1.1-4.1.5.

$$\begin{aligned}\mathbb{P}(X_n < k_1 \mid Z_n > k_2) &= \mathbb{E}\left[\frac{\sum_{1\le i\neq j\le Z_{k_1}} Z^{(k_1)}_{n-k_1,i} Z^{(k_1)}_{n-k_1,j}}{Z_n(Z_n-1)} \;\Big|\; Z_n > k_2\right] \\ &= \mathbb{E}\left[\frac{\left(\sum_{i=1}^{Z_{k_1}} Z^{(k_1)}_{n-k_1,i}\right)^2 - \sum_{i=1}^{Z_{k_1}} (Z^{(k_1)}_{n-k_1,i})^2}{Z_n(Z_n-1)} \;\Big|\; Z_n > k_2\right].\end{aligned} \tag{5.39}$$

Recall that each of the individuals in generation n must come from a branching process initiated by one of the Z_{k_1} individuals in generation k_1. Analogous to (5.27), we have

$$\sum_{i=1}^{Z_{k_1}} Z^{(k_1)}_{n-k_1,i} = Z_n. \tag{5.40}$$

We also may observe the following as $n \to \infty$:

$$\begin{aligned}\mathbb{E}(Z_n \mid Z_n > 0) \text{ grows in the order } n \quad &\Rightarrow \quad \left\{\frac{1}{Z_n} \;\Big|\; Z_n > 1\right\} \to 0 \\ &\Rightarrow \quad \left\{\frac{1}{Z_n} \;\Big|\; Z_n > n u_2\right\} \to 0 \\ &\Rightarrow \quad \left\{1 - \frac{1}{Z_n} \;\Big|\; Z_n > n u_2\right\} \to 1 \\ &\Rightarrow \quad \left\{\frac{Z_n(Z_n-1)}{Z_n^2} \;\Big|\; Z_n > n u_2\right\} \to 1.\end{aligned}$$

This confirms that under the present conditioning, $\frac{1}{Z_n(Z_n-1)}$ may be safely replaced by $\frac{1}{Z_n^2}$. In light of (5.40) and the observation above, (5.39) may be rewritten as

$$\begin{aligned}\mathbb{P}(X_n < k_1 \mid Z_n > k_2) &= \mathbb{E}\left[\frac{Z_n^2 - \sum_{i=1}^{Z_{k_1}} (Z^{(k_1)}_{n-k_1,i})^2}{Z_n^2} \;\Big|\; Z_n > k_2\right] \\ &= \mathbb{E}\left[1 - \frac{\sum_{i=1}^{Z_{k_1}} (Z^{(k_1)}_{n-k_1,i})^2}{\left(\sum_{i=1}^{Z_{k_1}} Z^{(k_1)}_{n-k_1,i}\right)^2} \;\Big|\; Z_n > k_2\right].\end{aligned} \tag{5.41}$$

We want to show that the latter term in (5.39) converges to $G(u_1, u_2)$ as $n \to \infty$ and concurrently $\frac{k_1}{n} \to u_1$ and $\frac{k_2}{n} \to u_2$. In order to see this, (5.41) shows that we may equivalently prove that

$$\mathbb{E}\left[\frac{\sum_{i=1}^{Z_{k_1}} (Z^{(k_1)}_{n-k_1,i})^2}{\left(\sum_{i=1}^{Z_{k_1}} Z^{(k_1)}_{n-k_1,i}\right)^2} \;\Big|\; Z_n > k_2\right] \longrightarrow 1 - G(u_1, u_2) = \mathbb{E}(\phi(V)) \tag{5.42}$$

as $n \to \infty$, $\frac{k_1}{n} \to u_1$, and $\frac{k_2}{n} \to u_2$.
Note that there may be some individuals from generation k_1 who have "died out" before getting to generation n. To this end, let G_{k_1} be the set of all individuals alive in the generation k_1. Then let $J_{k_1} \doteq \{i \in G_{k_1} : Z^{(k_1)}_{n-k_1,i} > 0\}$, the set of individuals in generation k_1 still alive at generation n. Then

$$\frac{\sum_{i=1}^{Z_{k_1}} (Z^{(k_1)}_{n-k_1,i})^2}{\left(\sum_{i=1}^{Z_{k_1}} Z^{(k_1)}_{n-k_1,i}\right)^2} = \frac{\sum_{i \in J_{k_1}} \left(\frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)^2}{\left(\sum_{i \in J_{k_1}} \frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)^2}. \tag{5.43}$$

In order to consider the asymptotic behavior of the point processes given here, we seek to utilize Theorem 5.1.2. J_{k_1} is a continuous functional of the point process $\{V_n\}$ as defined in Theorem 5.1.2 and hence weak convergence obtained in the theorem would suffice to give limiting behavior in this case. As a result, we may utilize the Theorem 5.1.2 to obtain the desired convergence under conditioning on the event $\{Z_n > k_2\}$. More precisely,

$$\frac{\sum_{i \in J_{k_1}} \left(\frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)^2}{\left(\sum_{i \in J_{k_1}} \frac{Z^{(k_1)}_{n-k_1,i}}{n-k_1}\right)^2} \longrightarrow \frac{\sum_{i=1}^{N_{u_1}} \eta_i^2}{\left(\sum_{i=1}^{N_{u_1}} \eta_i\right)^2}, \tag{5.44}$$

in the distributional sense, as $n \to \infty$, $\frac{k_1}{n} \to u_1 \in (0,1)$ and $\frac{k_2}{n} \to u_2 > 0$.

$$\frac{\sum_{i=1}^{N_{u_1}} \eta_i^2}{\left(\sum_{i=1}^{N_{u_1}} \eta_i\right)^2} = (\phi(V)),$$

where V is the point process generated by $\{\eta_i\}_{i\geq 1}$ and N_{u_1} with $\{\eta_i\}_{i\geq 1}$ i.i.d. exponential random variables with parameter $\frac{\sigma^2}{2}$ and N_{u_1} independent of $\{\eta_i\}_{i\geq 1}$ with distribution $\mathbb{P}(N_{u_1} = k) = (1 - u_1)u_1^{k-1}$, $k \geq 1$, having intensity given by $e^{\frac{2u_2}{\sigma^2}}$. This finishes the proof.

□

The goal of this research has been to expand upon previously known results pertaining to the coalescence time for the critical branching process. The main results of Chapter 4 clarify the asymptotic behavior of the total coalescence time under multi-scale conditioning. As an extension of these results, we also considered more precise conditioning in Section 4.2. Several interesting observations regarding the nature of these results leads to an open question, inviting more fine-tuned scaling on the random variable Y_n representing the total coalescence time. Chapter 5 takes advantage of the theory of point processes to obtain analogous results to those of Section 4.1 for the random variable representing the pairwise coalescence time, X_n.

BIBLIOGRAPHY

[1] Athreya, K. B. “Coalescence in critical and sub-critical Galton-Watson branching processes”. *Journal of Applied Probability* **49** (2012), pp. 627–638.

[2] Athreya, K. B. & Hong, J. “Coalescence on supercritical Bellman-Harris branching processes”. *Taiwanese Journal of Mathematics* **22 (1)** (2018), pp. 245–261.

[3] Athreya, K. B. & Ney, P. E. *Branching processes*. Dover, 2004.

[4] Billingsley, P. *Probability and Measure*. John Wiley & Sons, 1995.

[5] Cardona-Tobon, N. & Palau, S. “Yaglom’s limit for critcal Galton-Watson processes in varying environment: A probabilistic approach”. *Bernoulli* **27 (3)** (2021), pp. 1643–1665.

[6] Champagnat, N. & Villemonais, D. “Exponential convergence to quasi-stationary distribution and Q-process”. *Probability Theory and Related Fields* **164** (2016), pp. 243–283.

[7] Collet, P., Martinez, S. & San Martin, J. *Quasi-stationary distributions: Markov chains, diffusions, and dynamical systems*. Springer, 2013.

[8] Freitas, A., Freitas, J. & Magalhaes, M. “Complete covergence and recordsfor dynamically generated random processes”. *Transactions of the American Mathematical Society* (2017), p. 373.

[9] Grimmett, G. & Stirzaker, D. *Probability and Random Processes*. Oxford University Press, 2001.

[10] Grosjean, N. & Huillet, T. “On the genealogy and coalescence times of Bienayme-Galton-Watson branching processes”. *Stochastic Models* **34** (2018), pp. 1–24.

[11] Harris, T. E. *The Theory of Branching Processes*. Springer-Verlag Berlin Heidelberg, 1963.

[12] Jacod, J. & Protter, P. *Probability Essentials*. Springer, 2004.

[13] Jagers, P. “Galton-Watson processes in varying environments”. *Journal of Applied Probability* **11** (1974), pp. 174–178.

[14] Kallenberg, O. *Random Measures, Theory, and Applications*. Springer, 2017.

[15] Kolmogorov, A. “Zur Losung einer biologischen Aufgabe”. *Izvestiya nauchno - issledovatelskogo instituta matematiki i mechaniki pri Tomskom Gosudarstvennom Universitete* **2** (1938), pp. 1–6.

[16] Liu, J. “A Yaglom type asymptotic result for subcritical branching Brownian motion with absorption”. *Stochastic Processes and their Applications* **141** (2021), pp. 245–273.

[17] Vatutin, V. & Liu, Q. “Limit theorems for decomposable branching processes in a random environment”. *Journal of Applied Probability* **52** (2015), pp. 877–893.

[18] Yaglom, A. M. “Certain limit theorems of the theory of branching processes”. *Doklady Akademii Nauk SSSR* **56** (1947), pp. 795–798.