EXTINCTION IN SINGLE AND MULTI-TYPE BRANCHING PROCESSES

TIMOTHY CSERNICA

Abstract. Branching processes model a wide range of phenomena. The problem of extinction is of fundamental importance to branching processes. We describe both single and multi-type branching processes as sequences and construct probability generating functions for each term of these sequences. We then show that the probability of extinction is equal to the smallest nonnegative root of the fixed-point equation for the first generation's probability generating function. In both cases, we establish the conditions which determine this extinction probability.

Contents

1. Introduction

A *branching process* is a random process which proceeds through generations, each of which has some number of individuals. Every individual in generation n produces individuals in generation $n+1$ according to some distribution.

Branching processes have been used to model phenomena such as nuclear reactions, population growth, and cosmic rays [2]. One of the first (and most famous) applications is that of family names; the original problem was posed by Francis Galton in 1873 [1]:

> A large nation, of whom we will only concern ourselves with the adult males, N in number, and who each bear separate surnames, colonise a district. Their law of population is such that, in each generation, a_0 per cent of the adult males have no male children who reach adult life; a_1 have one such male child; a_2 have two; and so on up to a_5 who have five. [Throughout the paper, we will use p_k to refer to these probabilities.]

> Find (1) what proportion of the surnames will have become extinct after r generations; and (2) how many instances there will be of the same surname being held by m persons.

An answer was first presented by Reverend H. W. Watson [3]. We will first examine in detail the solution to question (1). We will then address the same question for a *multi-type* branching process.

A multi-type branching process allows for different kinds of objects in each generation. For example, in the context of Galton's problem, there may be type A males, who are especially fertile, and type B males, who are less fertile. Furthermore, each type will have its own offspring distribution; e.g., especially fertile males may be more likely to produce fertile offspring. We will examine the extinction problem posed above for a general multi-type branching process.

2. Mathematical Description of Single-type Branching Processes

We begin by defining a sequence which gives the number of individuals in each generation of the branching process.

Definition 2.1. For each generation n, we define Z_n as the number of individuals in the generation. We write (Z_n) for the sequence of the Z_n.

Remark 2.2. When dealing with single-type processes, we will assume that the process begins in the 0^{th} generation, and that $Z_0 = 1$.

Definition 2.3. (Offspring Distribution) Let $P(Z_1 = k) = p_k$. As we assumed $Z_0 = 1$, this means that the probability a single individual has k children is equal to p_k. Note that $\sum_{k=0}^{\infty} p_k = 1$. The p_ks define the *offspring distribution* of Z_0.

Remark 2.4. The distribution above is defined for a single individual. Z_{n+1} is the sum of Z_n independent random variables, all of which have the same distribution as Z_1.

It is expedient to use a probability generating function for the p_k.

Definition 2.5. A probability generating function (PGF) of a discrete random variable is a power series with coefficients equal to the values of the random variable's probability mass function. We can construct a probability generating function for Z_n where the coefficients are equal to $P(Z_n = k)$. In particular, the PGF for Z_1 is

$$f(s) = \sum_{k=0}^{\infty} p_k s^k,$$

where s is a complex variable with $|s| \leq 1$.

Remark 2.6. A PGF converges whenever $|s| \leq 1$. (If $|s| = 1$, then $\sum_{k=0}^{\infty} p_k = 1$, so the series converges absolutely. If $|s| < 1$, then $f(s)$ converges absolutely by comparison to a geometric series.) This implies that any PGF is smooth on the open unit disc.

Remark 2.7. For the rest of the paper, we will make the following assumptions:

(1) For all k, $p_k \neq 1$.

(2) $p_0 + p_1 < 1$.

(3) The expected value $E[Z_1] = \sum_{k=0}^{\infty} kp_k$ is finite. In particular, this implies that $f'(1)$ is finite.

We next define the *iterates* of a probability generating function, which allow us to express these functions in terms of f.

Definition 2.8. The iterates of the generating function $f(s)$ are defined by

$$f_0(s) = s$$

$$f_{n+1}(s) = f[f_n(s)].$$

Proposition 2.9. *For all $m, n \in \mathbb{N}$, we have $f_{m+n}(s) = f_m[f_n(s)]$.*

Proof.

$$f_{m+n} = f[f_{m+n-1}(s)] = f[f[f_{m+n-2}(s) = f_2[f_{m+n-2}(s)] = \cdots = f_m[f_n(s)].$$

□

A special case of the above proposition states that $f_{n+1}(s) = f_n[f(s)]$, an equality which we will soon make use of. Using a PGF and its iterates, we are able to derive the generating function of Z_n.

Theorem 2.10. [2]. *The generating function of Z_n is the n^{th} iterate $f_n(s)$.*

Proof. Let $f_{(n)}$ be the generating function of Z_n, $n = 0, 1, \ldots$. Let $k = Z_n$. The probability that generation Z_{n+1} has ℓ members is equal to the sum of the probabilities of all possible ways generation Z_n can have ℓ offspring. Let ℓ_i be the number of children of the i^{th} individual of generation Z_n. We have

$$P(Z_{n+1} = \ell \mid Z_n = k) = \sum_{\substack{(\ell_1,\ldots,\ell_k) \\ \ell_1+\cdots+\ell_k=\ell}} p_{\ell 1} p_{\ell 2} \cdots p_{\ell k} \tag{1}$$

where the sum is over all $(\ell_1, \ldots, \ell_k)$ such that $\ell_1 + \cdots + \ell_k = \ell$. (That is, all ways generation Z_n can have ℓ offspring.) Note that this holds because every individual reproduces independently.

By the definition of PGF, we know that

$$f_{(n+1)}(s) = \sum_{\ell=0}^{\infty} P(Z_{n+1} = \ell) s^{\ell}.$$

Using the Law of Total Probability, we have

$$f_{(n+1)}(s) = \sum_{\ell=0}^{\infty} s^{\ell} \sum_{k=0}^{\infty} P(Z_n = k) P(Z_{n+1} = \ell \mid Z_n = k).$$

We then apply equation (1):

$$f_{(n+1)}(s) = \sum_{k=0}^{\infty} P(Z_n = k) \sum_{\ell=0}^{\infty} s^{\ell} \sum_{\substack{(\ell_1,\ldots,\ell_k) \\ \ell_1+\cdots+\ell_k=\ell}} p_{\ell 1} p_{\ell 2} \cdots p_{\ell k}.$$

But the coefficients of $(f(s))^k$ are precisely the quantity in equation (1). Thus, we have

$$\begin{aligned} f_{(n+1)}(s) &= \sum_{k=0}^{\infty} P(Z_n = k)(f(s))^k \\ &= f_{(n)}[f(s)]. \end{aligned}$$

We also know that

$$f_{(0)}(s) = \sum_{k=0}^{\infty} P(Z_0 = k)s^k = s = f_0(s).$$

By induction, we may show $f_{(n)}(s) = f_n(s)$. □

Proposition 2.11. *For all $n \in \mathbb{N}$, we have $P(Z_n = 0) = f_n(0)$.*

Proof.

$$f_n(0) = \sum_{k=0}^{\infty} P(Z_n = k)0^k = P(Z_n = 0).$$

□

A special case of this proposition yields that $p_0 = f(0)$.

3. Extinction Problem for Single-type Branching Processes

Definition 3.1. (Extinction) A branching process becomes *extinct* if there exists some $N \in \mathbb{N}$ for which, if $n \geq N$, then $Z_n = 0$.

Definition 3.2. (Mean Number of Offspring) We write m for the mean number of offspring of a single individual. Because

$$f'(1) = \sum_{k=1}^{\infty} kp_k = 1p_1 + 2p_2 + 3p_3 + \cdots = m,$$

we know that $f'(1) = m$. From Remark 2.7, we know that m is finite.

Notation 3.3. Let q be the probability that the process goes extinct. We write q_n for the probability that the process goes extinct by generation n. Note that $q_0 = 0$, by the assumption that $Z_0 = 1$. Also note

$$0 \leq q_0 \leq q_1 \leq \cdots \leq q_n \leq 1.$$

Furthermore, by elementary measure theory,

$$\lim_{n \to \infty} q_n = q.$$

Lemma 3.4. [3, 4]. *Let $f(s)$ be the probability generating function of a branching process. The probability of extinction is equal to the smallest root of the fixed-point equation, $s = f(s)$.*

Proof. We will proceed by finding an expression for q_n. Let k equal the number of offspring in generation 1. We can consider each of these k offspring as the 0^{th} generation of a new branching process. We then have that q_n is the probability that the processes started by each of these k offspring are now extinct. Thus, we have that

$$q_n = \sum_{k=0}^{\infty} p_k (q_{n-1})^k.$$

This is precisely the generating function $f(q_{n-1})$, so $q_n = f(q_{n-1})$. Taking the limit yields that $q = f(q)$.
It remains to show that q is the smallest root. Because $P(Z_n = 0) = f_n(0)$,

$$q = \lim_{n\to\infty} q_n = \lim_{n\to\infty} f_n(0).$$

Probability generating functions are nondecreasing; thus, if there were another root, q^*, of the fixed point equation, we would have

$$q = \lim_{n\to\infty} f_n(0) \le \lim_{n\to\infty} f_n(q^*) = q^*.$$

Thus, q is the smallest root of the fixed-point equation. □

Lemma 3.5. [3]. *Let $s \in (0, 1]$. Both $f'(s)$ and $f''(s)$ are positive.*

Proof. We proceed first by differentiating f termwise. Recall that

$$f(s) = \sum_{k=0}^{\infty} p_k s^k = p_0 + p_1 s + p_2 s^2 + \cdots$$

so we have

$$f'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} = p_1 + 2p_2 s + 3p_3 s^2 + \cdots$$

and

$$f''(s) = \sum_{k=2}^{\infty} k(k-1) p_k s^{k-2} = 2p_2 + 3 \cdot 2p_3 s + 4 \cdot 3p_4 s^2 + \cdots$$

We know that $0 < s \le 1$. Additionally, all p_i are nonnegative, and from Remark 2.7, we know that $p_i > 0$ for some $i \ge 2$. Thus, we know $f'(s)$ and $''f(s)$ are positive. □

Theorem 3.6. [4]. *If $m \le 1$, then the process becomes extinct with probability 1. If $m > 1$, then the process becomes extinct with probability < 1.*

Proof. We must find the smallest root of the fixed-point equation. It is easy to verify that $s = 1$ is a solution, because if $s = 1$:

$$f(s) = f(1) = \sum_{k=0}^{\infty} p_k (1)^k = \sum_{k=0}^{\infty} p_k = 1,$$

as f is a probability generating function. Recall that m is the mean number of offspring, and that $m = f'(1)$. We will consider both the case $m \le 1$ and $m > 1$.
Case $\mathbf{1 : m \le 1}$
First, note that if $s = 0$ is a root, then $f(0) = p_0 = 0$. However, if $m \le 1$, then p_0 cannot equal 0, as this would violate the assumptions given in Remark 2.7. Thus, $s = 0$ is not a root. Now, assume for the sake of contradiction that some $s \in (0, 1)$ is a root, with $s = f(s)$. Because we know that $f''(s)$ is positive for all $s \in (0, 1]$, we have that for all $s \in (0, 1)$, the inequality $f'(s) < f'(1) = m \le 1$ holds. By the Mean Value Theorem, we have that

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

for some $c \in (s, 1)$. However, as $c \in (s, 1)$, we know that $f'(c) < 1$, so this is a contradiction. Thus, 1 is the smallest nonnegative root, so the probability of extinction is 1.

Case 2 : m > 1
Now, consider the case where $m > 1$. If $p_0 = 0$, we have that $p_0 = f(0) = 0$, giving the root $s = 0$, clearly the smallest root on $[0, 1]$. Suppose now that $p_0 > 0$, and note that $f(0) = p_0 > 0$. Also, as $f'(1) = m > 1$, by Taylor's Theorem, we have that $f(s) < s$ when s is sufficiently close to 1. Thus, by the Intermediate Value Theorem, there exists some $x \in (0, 1)$ such that $f(x) = x$. Therefore, whenever $m > 1$, there is a root on the interval $[0, 1)$, so the probability of extinction is less than 1. □

4. Mathematical Description of Multi-type Branching Processes

In many cases, the individuals involved in a branching process are not all alike. Examples include [5]:

(1): Population Genetics. When considering inheritance of alleles, a 3-type branching process, with types corresponding to the genotypes AA, Aa, and aa can be used as a model.

(2): Physics. Cosmic-ray cascades involve both electrons and photons, with electrons producing photons and photons producing electrons. Such an example is modeled by a 2-type branching process.

We begin by describing the offspring distributions for all individuals. Much of the content in sections 4-7 comes from [5]. We will always be dealing with a branching process containing m types of individuals.

Definition 4.1. For each generation n, we define $\mathbf{Z}_n$ as an m-dimensional vector whose i^{th} entry gives the number of individuals of type i in the n^{th} generation of the branching process. We write $(\mathbf{Z}_n)$ for the sequence of the $\mathbf{Z}_n$.

Remark 4.2. Unless otherwise noted, we will assume that $\mathbf{Z}_0 = \mathbf{e}_i$ for $i = 1, 2, \ldots, m$, where $\mathbf{e}_i$ is the i^{th} m-dimensional standard basis vector

Definition 4.3. (Offspring Vector) Consider m-dimensional vectors $\mathbf{k} = (k_1, k_2, \cdots, k_m)$, where $k_1, \ldots, k_m$ are natural numbers. These vectors represent the offspring of each type created by a individual, with k_1 equalling the number of offspring of type 1, and so on. We call the vector $\mathbf{k}$ the *Offspring Vector* for the individual.

Definition 4.4. (Offspring Distribution) For every type $i = 1, 2, \ldots, m$, let $p_i(\mathbf{k})$ be the probability that an individual of type i has the offspring vector $\mathbf{k}$. Note that $\sum_{\mathbf{k}} p_i(\mathbf{k}) = 1$. These probabilities define the *Offspring Distribution* for individuals of type i.

For single-type branching processes, we used the notation p_k to denote the probability an individual had k children. We now use the notation $p_i(\mathbf{k})$ to denote the probability that an individual of type i has Offspring Vector $\mathbf{k}$.

We next construct PGFs for a multi-type branching process.

Definition 4.5. (Probability Generating Function) Let $\mathbf{s}$ denote an m-dimensional vector $(s_1, s_2, \ldots, s_m)$ of complex numbers, and let $\max|s_i| = ||s||$. For all $i = 1, 2, \ldots m$ and all such vectors with $||s|| \leq 1$, let

$$f_i(\mathbf{s}) = \sum_{\mathbf{k}} p_i(\mathbf{k}) s_1^{k_1} s_2^{k_2} \cdots s_m^{k_m}.$$

We say f_i is the generating function for the offspring distribution of individuals of type i.

5. Multi-type Branching Processes as a Markov Chain

Construction 5.1. Let $\mathbf{N}_i$ with $i = 1, 2, \ldots, m$ be a vector-valued random variable defined such that $P[\mathbf{N}_i = \mathbf{k}] = p_i(\mathbf{k})$. Fix i and let $\mathbf{Y}_{ij}$ be a sequence of vector-valued random variables with the same distribution as $\mathbf{N}_i$.
Consider $\mathbf{Z}_n$. If $\mathbf{Z}_{n-1} = (Z_{1(n-1)}, Z_{2(n-1)}, \ldots, Z_{m(n-1)})$, we have

$$\mathbf{Z}_n = \sum_{i=1}^{m} \sum_{j=1}^{Z_{i(n-1)}} \mathbf{Y}_{ij}, \tag{2}$$

where if $Z_{i(n-1)} = 0$ for some i, the corresponding sum is interpreted as the zero vector. The above sum simply creates an appropriate offspring vector $\mathbf{k}$ for each individual in $\mathbf{Z}_{n-1}$ and adds them to yield $\mathbf{Z}_n$.
We now construct a generating function for the sequence $(\mathbf{Z}_n)$ by

$$F_n(i, \mathbf{s}) = \sum_{\mathbf{k}} P[\mathbf{Z}_n = \mathbf{k} \mid \mathbf{Z}_0 = \mathbf{e}_i] s_1^{k_1} s_2^{k_2} \cdots s_m^{km},$$

where $i = 1, 2, \ldots, m$ and $\mathbf{s} = (s_1, \ldots, s_m)$ with $\|s\| \leq 1$. Note both that

$$F_0(i, \mathbf{s}) = s_i$$

and

$$F_1(i, \mathbf{s}) = f_i(\mathbf{s}).$$

We find the other iterates of F by defining an m-dimensional vector of generating functions $\mathbf{f}$ by setting

$$\mathbf{f}(1, \mathbf{s}) = (f_1(\mathbf{s}), \ldots, f_m(\mathbf{s}))$$

and

$$\mathbf{f}(n, \mathbf{s}) = (f_1(\mathbf{f}(n-1, \mathbf{s})), \ldots, f_m(\mathbf{f}(n-1, \mathbf{s})))$$

for all $n \geq 2$. It then follows from equation (2) that

$$F_n(i, \mathbf{s}) = f_i(\mathbf{f}(n-1, \mathbf{s})) \tag{3}$$

for $n \geq 2$ and $i = 1, 2, \ldots, m$.
Additionally, using equation (2), we have that if $\mathbf{Z}_{n-1} = \mathbf{y}$, then the generating function of $\mathbf{Z}_n$ is

$$\prod_{i=1}^{m} f_i^{y_i}(\mathbf{s}). \tag{4}$$

Construction 5.2. Because all individuals reproduce independently, and offspring distributions do not change between generations, we know that $\mathbf{Z}_n$ depends only on $\mathbf{Z}_{n-1}$. Thus, $(\mathbf{Z}_n)$ is a Markov chain with stationary transition probabilities. For this chain, we define the state space X as the set of all m-dimensional vectors of natural numbers. For any pair of vectors $\mathbf{x}, \mathbf{y} \in X$, let $p(\mathbf{x}, \mathbf{y})$ be the coefficient of $s_1^{y_1} s_2^{y_2} \cdots s_m^{y_m}$ in the generating function given in equation (4). We have that

$$P[\mathbf{Z}_n = \mathbf{y} \mid \mathbf{Z}_{n-1} = \mathbf{x}] = p(\mathbf{x}, \mathbf{y}) \tag{5}$$

for all $n \geq 1$, and all vectors $\mathbf{x}, \mathbf{y} \in X$, meaning that $(p(\mathbf{x}, \mathbf{y}))$ is the matrix of transition probabilities.

We now use the fact that $(\mathbf{Z}_n)$ is a Markov chain to prove an important result about the recurrence of states.

6. Positively Regular Processes and Irreducible States

The nonreccurence of states in Multi-type Branching processes is necessary to prove extinction theorems. We begin by what it means for a process to be positively regular:

Definition 6.1. (Positively Regular) A multi-type process is irreducible if and only if, for every pair of types i, j, there exists some natural number n such that

$$P[Z_{jn} \geq 1 \mid \mathbf{Z}_0 = \mathbf{e}_i] > 0.$$

Here, Z_{jn} gives the number of individuals of the jth type in generation n. If, for some n, the statement holds for all i, j, then the process is *positively regular*.

Construction 6.2. (Associated Matrix) For any multi-type branching process, there is an associated $m \times m$ matrix $\mathbf{M}$ defined by

$$m_{ij} = E[Z_{j1} \mid \mathbf{Z}_0 = \mathbf{e}_i]$$

for $i, j = 1, 2, \ldots, m$. We can see the relationship between $\mathbf{M}$ and $\mathbf{Z}_n$ by the following construction: Let

$$\mathbf{u}_n = E\mathbf{Z}_n = (u_{1n}, \ldots, u_{mn}).$$

Recall equation (2):

$$\mathbf{Z}_n = \sum_{i=1}^{m} \sum_{j=1}^{Z_{i(n-1)}} \mathbf{Y}_{ij}.$$

Taking the expected value of (2) yields that

$$\mathbf{u}_n = \mathbf{u}_{n-1}\mathbf{M}$$

and therefore

$$\mathbf{u}_n = \mathbf{u}_0\mathbf{M}^n. \tag{6}$$

If a multi-type branching process is positively regular, there exists a natural number n such that all elements in $\mathbf{M}^n$ are strictly positive. We now define what it means for the matrix $\mathbf{M}$ to be positively regular, which is stricter than the process being positively regular.

Definition 6.3. (Positively regular matrix) The matrix $\mathbf{M}$ is positively regular if all elements of $\mathbf{M}$ are finite and there exists a natural number n such that all elements of $\mathbf{M}^n$ are strictly positive.

We will need a few properties of positively regular matrices to deal with extinction:

Theorem 6.4. *(Perron–Frobenius Theorem) If $\mathbf{M}$ is a positively regular matrix, then $\mathbf{M}$ has a positive eigenvalue r (the* Perron–Frobenius *root of $\mathbf{M}$) of multiplicity 1, such that for all other eigenvalues λ of $\mathbf{M}$, we have that $|\lambda| < r$. The eigenvalue r has associated right and left eigenvectors $\mathbf{v}$ and $\mathbf{w}$ with strictly positive elements such that $r\mathbf{v} = \mathbf{v}\mathbf{M}$ and $\mathbf{M}\mathbf{w} = r\mathbf{w}$, and $\mathbf{v}\mathbf{w} = 1$. Let $\mathbf{M}_1 = \mathbf{w}\mathbf{v} = (\mathbf{w}_i v_j)$ and $\mathbf{M}_2 = \mathbf{M} - r\mathbf{M}_1 = (m_{2ij})$. It follows that $\mathbf{M}_1\mathbf{M}_1 = \mathbf{M}_1$, $\mathbf{M}_1\mathbf{M}_2 = \mathbf{M}_2\mathbf{M}_1 = \mathbf{0}$, and*

$$\mathbf{M}^n = r^n\mathbf{M}_1 + \mathbf{M}_2^n \tag{7}$$

for all $n \in \mathbb{N}$. Furthermore, there exists a positive number α with $0 < \alpha < r$ such that

$$m_{2ij}^{(n)} = O(\alpha^n) \tag{8}$$

where $\mathbf{M}_2^n = (m_{2ij}^{(n)})$. The term $O(\alpha^n)$ states that some constant times α^n is an upper bound for $m_{2ij}^{(n)}$ as n approaches infinity. This bound is independent of i, j.

Proofs of the previous statements may be found in [6, Appendix 2].

We now move into nonrecurrence.

Definition 6.5. (Nonreccurent State) A state $\mathbf{k} \in X$ is nonreccurent if and only if, for some $t \geq 1$,

$$P[\mathbf{Z}_{n+t} = \mathbf{k} \mid \mathbf{Z}_n = \mathbf{k}] < 1.$$

Remark 6.6. If a Markov Chain has stationary transition probabilities, the following definition is equivalent: A state is nonreccurent if and only if

$$P[\mathbf{Z}_{n+t} = \mathbf{k} \text{ infinitely often } \mid \mathbf{Z}_n = \mathbf{k}] = 0.$$

The following definition allows us to avoid the case where each individual produces one offspring with probability one:

Definition 6.7. A multi-type branching process is *singular* if the generating functions $f_1(\mathbf{s}), \dots f_m(\mathbf{s})$ are all linear in $\mathbf{s}$ with no constant terms; i.e., each object has exactly one child.

We will now prove that, in a nonsingular positively regular process, all states other than $\mathbf{0}$ are nonrecurrent.

Definition 6.8. [2]. Let S be a set of types i such that

$$\text{if } P(\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i) = 0 \text{ for } n = 1, 2, \dots, \text{ then } i \in S.$$

We assume that the types in S are numbered $1, 2, \dots, r$. We may assume this without loss of generality through relabelling.

Lemma 6.9. [2]. *If S is not empty, then*

$$P(Z_{1(n+1)} + \cdots + Z_{r(n+1)} \geq Z_{1n} + \cdots + Z_{rn}) = 1, n = 0, 1, \dots$$

Proof. Let $\mathbf{k} \in X$ with $k_1 + \cdots + k_r = 0$. By the definition of S and equation (6), we have that

$$P(\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{k}) > 0$$

when n is sufficiently large. It follows that

$$P(\mathbf{Z}_1 = \mathbf{k} \mid \mathbf{Z}_0 = \mathbf{e}_i) = 0, i \in S.$$

As this holds for all $\mathbf{k}$ with $k_1 + \cdots + k_r = 0$, we know that

$$P(Z_{11} + \cdots + Z_{r1} \geq 1 \mid \mathbf{Z}_0 = \mathbf{e}_i) = 1, i \in S.$$

Because we assumed $\mathbf{Z}_0 = \mathbf{e}_i$ for some i, we have that

$$P(Z_{11} + \cdots + Z_{r1} \geq Z_{10} + \cdots + Z_{r0}) = 1.$$

The lemma then follows because the process satisfies the Markov property. □

Theorem 6.10. [2]. *If a branching process is positively regular and nonsingular, then all nonzero states are nonrecurrent.*

Proof. Fix some nonnegative $\mathbf{k} \in X$. We will consider two cases:
Case 1 : S is empty.
For each i, the probability $P(\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i) > 0$ for all sufficiently large n. Any process beginning with $\mathbf{Z}_0 = \mathbf{k}$ may be thought of as a combination of these processes; thus we have that $P(\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{k}) > 0$ for sufficiently large n. Therefore, $\mathbf{k}$ is nonrecurrent.
Case 2 : S is not empty.
We will begin by assuming that $\mathbf{M}$ is positive.
For any vector $\mathbf{w}$, let $\pi(\mathbf{w})$ be the sum of the first r components of $\mathbf{w}$. We will show that for some $b \in \mathbb{N}$, we have

$$P(\pi(\mathbf{Z}_b) > \pi(\mathbf{k}) \mid \mathbf{Z}_0 = \mathbf{k}) > 0. \tag{9}$$

Now, if $n > b$, then Lemma 6.9 implies that the conditional probability that $\mathbf{Z}_n = \mathbf{k}$, given that $\pi(\mathbf{Z}_b) > \pi(\mathbf{k})$, is 0. Thus, the state $\mathbf{k}$ is nonrecurrent.

We will first suppose that $r = m$. In this case, because the process is nonsingular, for some $i = 1, 2, \ldots, m$, we have $P(\pi(\mathbf{Z}_1) > 1 \mid \mathbf{Z}_0 = \mathbf{e}_i) > 0$. Let j be such a value of i. If $\mathbf{w} \in X$ is such that $w_j \geq 1$, then by the definition of the process, $P(\pi(\mathbf{Z}_2) > \pi(\mathbf{w}) \mid \mathbf{Z}_1 = \mathbf{w}) > 0$. Because $\mathbf{M}$ is positive, there is a positive probability that $Z_{j1} \geq 1$ when $\mathbf{Z}_0 = \mathbf{k}$. Furthermore, from Lemma 6.9, we have that $P(\pi(\mathbf{Z}_2) \geq \pi(\mathbf{Z}_1)) = 1$. Therefore, if $r = m$, then equation (9) holds for $b = 2$.

Consider now when $1 \leq r < m$. Let $\mathbf{Z}_0 = \mathbf{k}$. Given some $\mathbf{w} \in X$ with $\mathbf{Z}_1 = \mathbf{w}$, we have that $\mathbf{Z}_2$ is the sum of $\mathbf{Z}_2'$ and $\mathbf{Z}_2''$, the children of the first r and last $k - r$ components of $\mathbf{w}$, respectively. Using Lemma 6.9, we have that $\pi(\mathbf{Z}_2') \geq \pi(\mathbf{w})$ with probability 1. For the same reason, we know $\pi(\mathbf{w}) \geq \pi(\mathbf{k})$. Because $\mathbf{M}$ is positive,we know that $Z_{(r+1)1}$ has a positive probability of being positive, which implies that $\pi(\mathbf{Z}_2'')$ has a positive probability of being positive. Because $\pi(\mathbf{Z}_2) = \pi(\mathbf{Z}_2') + \pi(\mathbf{Z}_2'')$, equation (9) holds with $b = 2$. Thus, the theorem holds when $\mathbf{M}$ is positive.

Now, consider when $\mathbf{M}^n$ is positive for some n. Recall equation (6):

$$E\mathbf{Z}_n = \mathbf{Z}_0\mathbf{M}^n.$$

We may obtain that

$$E(\mathbf{Z}_{n+N} \mid \mathbf{Z}_N) = \mathbf{Z}_N\mathbf{M}^n, n, N = 0, 1, \ldots$$

Using this equation, we have that for a fixed $\mathbf{Z}_N$, the conditional distribution of $\mathbf{Z}_{n+N}$ is the sum of $Z_{1N} + \cdots + Z_{mN}$ independent random vectors, of which $\mathbf{Z}_{iN}$ have the generating function $f_{in}, i = 1, 2, \ldots, m$. We therefore have that the vectors $\mathbf{Z}_0, \mathbf{Z}_N, \mathbf{Z}_{2N}, \ldots$ are a branching process (the 'N'-process) with associated matrix $\mathbf{M}^N$.

We will now show the 'N' process is nonsingular. Assume that it were singular. This would imply that $F_N(i, \mathbf{0}) = 0$ for $i = 1, 2, \ldots, m$. Because $F_n(i, \mathbf{0})$ is nondecreasing in n, this implies that $f_i(\mathbf{s}) = 0$ for $i = 1, 2, \ldots, m$. As the original process is nonsingular, some f_i must be nonlinear. Assume that it is f_1. We then see that $F_2(1, \mathbf{s}) = f_1(\mathbf{f}(1, \mathbf{s}))$ is nonlinear, because the f_i have no constant terms. We may show by induction that $F_n(1, \mathbf{s})$ is nonlinear for all n, contradicting the assumption that the original process is singular.

Note that the set S is the same for the 'N' process as for the original process. Because the associated matrix of the 'N' process is positive, the argument for a branching process where $\mathbf{M}$ is positive applies. Thus, we see that equation (9) holds for some value of b. This implies the truth of the theorem for the original process, showing the theorem holds in all cases. □

We now have the necessary tools to evaluate extinction probability.

7. Extinction Problem for Multi-type Branching Processes

Definition 7.1. A multi-type branching process becomes *extinct* if there exists some $N \in \mathbb{N}$ for which, if $n \geq N$, then $\mathbf{Z}_n = \mathbf{0}$, where $\mathbf{0}$ is the m-dimensional zero vector.

Notation 7.2. Recall that we assumed that $\mathbf{Z}_0 = \mathbf{e}_i$ for $i = 1, 2, \ldots, m$, where $\mathbf{e}_i$ is the i^{th} m-dimensional standard basis vector. We write q_i for the probability the process goes extinct when $\mathbf{Z}_0 = \mathbf{e}_i$. We then use q_{in} as the probability that the process beginning with $\mathbf{Z}_0 = \mathbf{e}_i$ goes extinct by generation n. We have that

$$q_{i0} \leq q_{i1} \leq \cdots \leq q_{in}$$

and, using some elementary measure theory,

$$q_i = \lim_{n\to\infty} q_{in}$$

for $i = 1, 2, \ldots, m$. We let $\mathbf{q}_n = (q_{1n}, \cdots, q_{mn})$ be the m-dimensional vector of the extinction probabilities, and write

$$\mathbf{q} = \lim_{n\to\infty} \mathbf{q}_n.$$

Lemma 7.3. *All q_i satisfy the equation $q_i = f_i(\mathbf{q})$.*

Proof. We will proceed by using the generating functions to identify the q_{in}. Note that

$$f_i(\mathbf{0}) = \sum_{\mathbf{k}} p_i(\mathbf{k}) 0^{k_1} \cdots 0^{k_n} = p_i(\mathbf{0}) = q_{i1}$$

Recall equation (3):

$$F_n(i, \mathbf{s}) = f_i(\mathbf{f}(n-1, \mathbf{s})).$$

From this equation, we have that

$$q_{in} = f_i(\mathbf{q}_{n-1}) \tag{10}$$

for $i = 1, 2, \ldots, m$. We know that the probabilities $\mathbf{q}_n$ are nondecreasing and bounded, and that generating functions are continuous in each variable. Thus, we may take the limit of equation (10), yielding

$$q_i = f_i(\mathbf{q})$$

for $i = 1, 2, \ldots, m$. □

We can make definite statements about solutions to this equation when the matrix $\mathbf{M}$ is positively regular. The rest of the paper is devoted to proving two theorems. The first states that the Perron–Frobenius root of $\mathbf{M}$ for a multi-type branching process plays a similar role to the mean number of offspring in a single-type process with regard to extinction. The second gives a method of calculating extinction probability, which uses $\mathbf{f}$, the vector of generating functions, and any vector in the unit cube except $\mathbf{1}$. It also shows that such a solution is unique.

Theorem 7.4. *Let $\mathbf{M}$ be a positively regular matrix and let r be the Perron–Frobenius root of $\mathbf{M}$. If $r \leq 1$, then $\mathbf{q} = \mathbf{1}$, the m-dimensional vector $(1, 1, \ldots, 1)$. If $r > 1$, then $q_i < 1$ for all $i = 1, 2, \ldots, m$. In either case, $\mathbf{q}$ is such that, for any other nonnegative solution $\mathbf{q}^*$ of the equation $\mathbf{q} = \mathbf{f}(\mathbf{q})$, we have that $q_i \leq q_i^*$ for all $i = 1, 2, \ldots, m$.*

Theorem 7.5. *Let the conditions of the above theorem hold. (i) For every m-dimensional nonnegative vector $\mathbf{s} = (s_1, \ldots, s_m)$ such that $||\mathbf{s}|| \leq 1$ but $\mathbf{s} \neq \mathbf{1}$, we have*

$$\lim_{n\to\infty} \mathbf{f}(n, \mathbf{s}) = \mathbf{q}.$$

(ii) Additionally, the only nonnegative solutions of the equation $\mathbf{f}(\mathbf{s}) = \mathbf{s}$ with $||\mathbf{s}|| \leq 1$ are $\mathbf{1}$ and $\mathbf{q}$.

Proof. (7.4) First, we will prove that if $r \leq 1$, then $\mathbf{q} = \mathbf{1}$. As $\mathbf{M}$ is a positively regular matrix, all states are nonrecurrent. Thus, for all $i = 1, 2, \ldots, m$ and $\mathbf{k} \in X \setminus \{\mathbf{0}\}$, we have

$$\lim_{n\to\infty} P[\mathbf{Z}_n = \mathbf{k} \mid \mathbf{Z}_0 = \mathbf{e}_i] = 0.$$

It follows that

$$P[\mathbf{Z}_n \to \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i] + P[\mathbf{Z}_n \to \infty \mid \mathbf{Z}_0 = \mathbf{e}_i] = 1.$$

Using equations (6) and (7) yields that

$$E\mathbf{Z}_n = \mathbf{u}_n = \mathbf{u}_0\mathbf{M}^n = \mathbf{u}_0 r^n \mathbf{M}_1 + \mathbf{M}_2^n.$$

As $r \leq 1$, the first term of this equation is bounded. From equation (8), we have that $\mathbf{M}_2^n$ is bounded as well. Therefore, $E[\mathbf{Z}_{jn} \mid \mathbf{Z}_0 = \mathbf{e}_i]$ is bounded for all $n \in \mathbb{N}$ and $i, j = 1, 2, \ldots, m$. Thus, we have that

$$P[\mathbf{Z}_n \to \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i] = 1,$$

so we know all states are nonrecurrent.

We will now show that if $r > 1$, then $q_i < 1$ for all $i = 1, 2, \ldots, m$.
Note that, if there exists an $N \in \mathbb{N}$ such that $f_i(N, \mathbf{0}) = 1$, then for $n \geq N$, all elements of the i^{th} row of $\mathbf{M}^n$ would be 0. If this were so, then $\mathbf{M}$ would not be positively regular; thus, we conclude that $f_i(n, \mathbf{0}) < 1$ for all $n \in \mathbb{N}$.
Also note that the column vector of functions $\mathbf{f}(n, \mathbf{1} - \mathbf{s})$ may be written as

$$\mathbf{f}(n, \mathbf{1} - \mathbf{s}) = \mathbf{1} - \mathbf{M}^n\mathbf{s} + o(\|\mathbf{s}\|) \tag{11}$$

where $o(\|\mathbf{s}\|)$ is a column vector such that for every ϵ there exists a δ such that $\|o(\|\mathbf{s}\|)\| \leq \epsilon\|\mathbf{s}\|$ for all $|\mathbf{s}| \leq \delta$. Because $r > 1$, for all nonnegative vectors $\mathbf{s} \neq \mathbf{0}$, there exists some $n \in \mathbb{N}$ such that

$$\|\mathbf{M}^n\mathbf{s}\| > 2\|\mathbf{s}\|.$$

From equation (11), we see that we can choose $\epsilon > 0$ such that if $\|\mathbf{s}\| < \delta$ then

$$\|\mathbf{1} - \mathbf{f}(n, \mathbf{1} - \mathbf{s})\| > \|\mathbf{s}\|.$$

Suppose that $\mathbf{q} = \mathbf{1}$. Then, there exists some $N \in \mathbb{N}$ such that $\|\mathbf{1} - \mathbf{f}(N, \mathbf{0})\| < \epsilon$. Because

$$\mathbf{f}(n + N, \mathbf{s}) = \mathbf{f}(n, \mathbf{f}(N, \mathbf{s})),$$

we know that

$$\|\mathbf{1} - \mathbf{f}(n+N, \mathbf{0})\| = \|\mathbf{1} - \mathbf{f}(n, \mathbf{f}(N, \mathbf{0}))\| > \|\mathbf{1} - \mathbf{f}(N, \mathbf{0})\|.$$

However, we also know that

$$f_i(n+N, \mathbf{0}) \geq f_i(N, \mathbf{0})$$

for $i = 1, 2, \ldots, m$, implying that

$$\|\mathbf{1} - \mathbf{f}(n+N, \mathbf{0})\| \leq \|\mathbf{1} - \mathbf{f}(N, \mathbf{0})\|.$$

This is clearly a contradiction, meaning that for some $i = 1, 2, \ldots, m$, we have $q_i < 1$.

It remains to show that $q_i < 1$ for all $i = 1, 2, \ldots, m$. Suppose that there exists some j such that $q_i = 1$ for $i = 1, 2, \ldots, j$ and $q_i < 1$ for $j+1, \ldots, m$. We may make this assumption without loss of generality because we can reorder the types. Let $n \in \mathbb{N}$ be such that all elements of $\mathbf{M}^n$ are positive. Recall that $q_i = f_i(\mathbf{q})$. Using this equation, we have that

$$1 = f_i(1, \ldots, 1, q_{j+1}, \ldots, q_m)$$

for $i = 1, 2, \ldots, j$. For nonnegative vectors, $f_i(n, \mathbf{s})$ is nondecreasing in each variable. Thus, the above equation implies that $f_i(n, \mathbf{s})$ is constant in s_i for $s_i \geq q_i$ and $i = j+1, \ldots, m$. Therefore, by differentiating $f_i(n, \mathbf{s})$ at the vector $\mathbf{1}$ we obtain that some $m_{ij}^{(n)}$ will be zero, contradicting the fact that $\mathbf{M}$ is positively regular. Thus, we conclude that $q_i < 1$ for all $i = 1, 2, \ldots, m$.

We will now show that $\mathbf{q}$ is the smallest nonnegative solution of $\mathbf{q} = f(\mathbf{q})$. Let $\mathbf{q}$ be a solution of $\mathbf{q} = f(\mathbf{q})$, and let $\mathbf{q}^*$ be any other nonnegative solution. Recall that $q_{i1} = f_i(\mathbf{0})$. Note that

$$F_n(i, \mathbf{0}) = \sum_{\mathbf{k}} P[\mathbf{Z}_n = \mathbf{k} \mid \mathbf{Z}_0 = \mathbf{e}_i] 0^{k_1} 0^{k_2} \cdots 0^{k_m} = P[\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i].$$

Therefore, we have

$$q_i = \lim_{n \to \infty} q_{in} = \lim_{n \to \infty} F_n(i, \mathbf{0}).$$

As generating functions are nondecreasing, for all $i = 1, 2, \ldots, m$, we have that

$$q_i = \lim_{n \to \infty} F_n(i, \mathbf{0}) \leq \lim_{n \to \infty} F_n(i, \mathbf{q}^*) = q_i^*.$$

□

Proof. (7.5) First, suppose that $\|\mathbf{s}\| < 1$. We know that

$$f_i(n, \mathbf{s}) = P[\mathbf{Z}_n = \mathbf{0} \mid \mathbf{Z}_0 = \mathbf{e}_i] + \sum_{\mathbf{k} \neq \mathbf{0}} P[\mathbf{Z}_n = \mathbf{k} \mid \mathbf{Z}_0 = \mathbf{e}_i] s_1^{k_1} s_2^{k_2} \cdots s_m^{k_m}$$

by the definition of f_i. Taking the limit as $n \to \infty$ yields that the first term approaches q_i, by definition, and that the second term approaches 0, as all $\mathbf{k} \neq \mathbf{0}$ are nonrecurrent. Using the Dominated Convergence Theorem gives that (i) is true for all $\|\mathbf{s}\| < 1$.

Now, suppose that $s_i = 1$ for some $i = 1, 2, \ldots, m$, but $s_j < 1$ for some $i \neq j$. We therefore know that for $n \in \mathbb{N}$ such that all components of $\mathbf{M}^n$ are positive, all of the components in $\mathbf{f}(n, \mathbf{s})$ are strictly less than 1. Because

$$\mathbf{f}(N, \mathbf{s}) = \mathbf{f}(N - n, \mathbf{f}(n, \mathbf{s}))$$

for all $N > n$, by definition, the argument used for the $\|\mathbf{s}\| < 1$ case may be applied here. Thus, (i) holds for all $\|\mathbf{s}\| \leq 1$.

We now prove (ii). Let $\mathbf{s}$ be any m-dimensional vector such that $\|\mathbf{s}\| \leq 1$ but $\mathbf{s} \neq \mathbf{1}$, and with

$$\mathbf{s} = \mathbf{f}(\mathbf{s}).$$

Iteration yields that

$$\mathbf{s} = \mathbf{f}(n, \mathbf{s})$$

for all $n \in \mathbb{N}$. Therefore, we have that

$$\mathbf{s} = \lim_{n \to \infty} \mathbf{f}(n, \mathbf{s}) = \mathbf{q},$$

which proves (ii). □

Acknowledgments. It is a pleasure to thank my mentor, Jon Sheperd, for his numerous comments, edits, and explanations, which were instrumental to the writing of this paper.

References

[1] D.G. Kendall. Branching Processes Since 1873. *Journal of London Mathematics Society.* vol. 41 (1966). p. 386.
[2] T. E. Harris. The Theory of Branching Processes. Springer-Verlag, Berlin. 1963.
[3] Charles Grinstead, J. Laurie Snell. Introduction to Probability.
[4] http://galton.uchicago.edu/ lalley/Courses/312/Branching.pdf
[5] Charles. J. Mode. Multitype Branching Processes: Theory and Applications. American Elsevier Publishing Company, Inc. New York. 1971.
[6] Samuel Karlin. A First Course in Stochastic Processes. Academic Press Inc. New York, New York. 1966.