THE TOTAL PROGENY OF A MULTITYPE BRANCHING PROCESS

THE TOTAL PROGENY OF A MULTITYPE BRANCHING PROCESS

By

Xingli Wei, M.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2008)	McMaster University
(Statistics)	Hamilton, Ontario

TITLE: The Total Progeny of a Multitype Branching Process

AUTHOR: Xingli Wei

SUPERVISOR: Dr. Fred M. Hoppe

NUMBER OF PAGES: vi, 42

Acknowledgements

I would like to thank my supervisor, Fred M. Hoppe, for his valuable comments and support throughout this thesis. In addition I would like to thank many people who have taught me mathematics at McMaster University, in particular McKenzie Wang, Shui Feng and Stanley Alama for their patiently teaching me the theory of linear algebra, probability theory and analysis, and for their kindly writing references for my graduate study. Finally I would like to thank my wife Hua and my daughter Cynthia for their support and patience during this not always easy year.

Contents

Abstract

Techniques from algebra and matrix theory are employed to study the total progeny of a multitype branching process from the point of probability generating functions. A result for the total progeny of different types of individuals having identical offspring distribution is developed, which extends the classic Dwass formula from single case to multitype case. An example with Poisson distributed offspring having different distributions of children is given to illustrate that total progeny does not preserve similar structure as Dwass' formula in general.

Chapter 1

Introduction

The purpose of this thesis is to extend to multitype branching processes the result of M. Dwass; he obtained a very beautiful formula for the total progeny for single type branching process with k initial ancestors in 1969, which is in general called Dwass' formula. I.J. Good obtained a formula for multitype (say d-type) branching process with initial $(k_1, k_2, \cdots, k_d)$ individuals in 1960, but his result, having different structure from Dwass', is not very useful computationally. Based on the work of Good, my result has similar structure to Dwass'. When reduced to single type case, it's consistent with Dwass'.

1.1 Background and Motivation

The proposed work is motivated by F.M. Hoppe. His work [2007] concerns the analysis of the effect of parlaying bets in lotteries. In that paper, a single type discrete-time branching process model for the effects has been presented, and the duration of parlay derived.

In Lotto Super 7, a single $2 ticket gives three sets of seven numbers each from 1 to 47. The player will win whenever at least three of the numbers in any of the sets of the player's ticket match some of the numbers drawn by the lottery. The parlaying occurs when a ticket wins a free ticket for a succeeding game or a small dollar prize that is used to purchase tickets for future games.

It is necessary to determine the distribution of the number of tickets available for each game subsequent to the first that are generated by parlaying. This corresponds to what is known as the total progeny. The appropriate tool for determining the distribution of the total progeny is a result of Dwass[1969], who gave a very nice formula of total progeny in a single type branching process with initial population size k.

Sometimes people maybe continue play the game when they win $10 or more. It is interesting to extend the result of Hoppe to multitype case. This motivated us to find a formula similar to Dwass' for multitype branching process. In the present thesis, we discuss this problem.

1.2 Overview of Branching Processes

Branching processes were introduced by Francis Galton in the nineteenth century as a simple mathematical model for the propagation of family names in British peerage. The subject of branching processes is well-developed, and has lead to successful applications in the areas of population dynamics, molecular biology, cell ecology, medicine, algorithms, combinatorics and others.

1.2.1 Branching Processes: Single Type

Let $\{Z_n\}_0^\infty$ be a branching process with initial population size Z_0 and offspring distribution $(p_i, i = 0, 1, \cdots)$. Here individuals may represent people, organisms, etc., depending on the context. We interpret Z_n as the number of individuals in the n^{th} generation. Each n^{th} generation individual produces a random number of individuals (called offspring) in the $(n+1)^{th}$ generation with identically and independent distribution $\{p_i\}_{i\geq 0}$.

1.2.2 Branching Process: Multitype

Definition 1.1. Let T denote the set of all d-dimensional vectors whose components are nonnegative integers. Let $\underline{e}_i$, $1 \leq i \leq k$, denote the vector whose i^{th} component is 1 and whose other components are 0.

Definition 1.2. The multitype (d-type) branching process is a temporally homogeneous vector Markov process $\{\underline{Z}_n\}, n = 0, 1, 2, \cdots$, whose states are vectors in T. We write

$$\underline{Z}_n = (Z_{n,1}, \cdots, Z_{n,d}) \tag{1.1}$$

where $\underline{Z}_{n,j}, j = 1, 2, \cdots, d$, denotes the number of type j objects in the n^{th} generation. If the process is initiated in state $\underline{i}$, we will be denote it by $\underline{Z}_n^{(\underline{i})}$.

Definition 1.3. If $\underline{Z}_0 = \underline{e}_i$, then $\underline{Z}_1$ have the generating function

$$f^{(i)}(\underline{s}) = \sum_{r_1,\cdots,r_d=0}^{\infty} p^{(i)}(r_1, \cdots, r_d)\, s_1^{r_1} \cdots s_d^{r_d} \tag{1.2}$$

where $p^{(i)}(r_1, \cdots, r_d)$ is the probability that an object of type i has r_1 children of type 1, $\cdots$, r_d children of type d. We sometimes write $f^{(i)}(\underline{s}) = f_1^{(i)}(\underline{s})$

If initially there are $i_1, \cdots, i_d$ individuals of types $1, \cdots, d$, respectively, that is, $\underline{Z}_0 = \underline{i}$, then the generating function for $\underline{Z}_1$ is

$$f(\underline{s}) = \left(f^{(1)}(\underline{s})\right)^{i_1} \cdots \left(f^{(d)}(\underline{s})\right)^{i_d} \tag{1.3}$$

The generating function of $\underline{Z}_n$, when $\underline{Z}_0 = \underline{e}_i$, will be denoted by

$$f_n^{(i)}(s_1, \cdots, s_d) = f_n^{(i)}(\underline{s}) \qquad n = 0, 1, \cdots; \ i = 1, \cdots, d \tag{1.4}$$

1.2.3 Total Progeny in a Branching Process

Definition 1.4. Let $\underline{S}$ be total progeny in a branching process $\{\underline{Z}_n\}$, then

$$\underline{S} = \underline{Z}_0 + \underline{Z}_1 + \cdots \tag{1.5}$$

Lemma 1.1. (*Good*[1955]) *Let there be d types, and let $w^{(i)}(\underline{s})$ be the generating function for the total numbers of the various types in all generations, starting with one object of type i. Then the $w^{(i)}(\underline{s})$ satisfy the functional equations*

$$w^{(i)}(\underline{s}) = s_i f^{(i)}\left(w^{(1)}(\underline{s}), \cdots, w^{(d)}(\underline{s})\right) \quad i = 1, \cdots, d$$

For d=1, this reduces to

$$w(s) = sf(w(s))$$

Lemma 1.2. *If the branching process starts with i_1 individuals of type 1, i_2 individuals of type 2,$\cdots$, i_d individuals of type d, then the generating function for the total numbers of the various types in all generations is given by*

$$w(\underline{s}) = (w^{(1)}(\underline{s}))^{i_1} \cdots (w^{(d)}(\underline{s}))^{i_d} \tag{1.6}$$

Theorem 1.3. *(Dwass[1969])*

$$\begin{aligned} P(S = m \mid Z_0 = i) &= \frac{i}{m}[s^{m-i}]f^m(s) \\ &= \frac{i}{m}P(Z_1 = m - i \mid Z_0 = m) \end{aligned} \tag{1.7}$$

or let $N_m = X_1 + \cdots + X_m$, *where* X_j *are iid. Then*

$$P(S = m) = \frac{i}{m}P(N_m = m - i) \tag{1.8}$$

The relationship (1.8) is called Dwass' formula.

The proof by Dwass is quite complicated. The technique he used involved the probability generating function of an infinitely divisible distribution. In fact, we can get his result by using Lagrange-Bürmann inversion formula. Harris[1963] mentioned that the distribution of the total progeny can be determined by using Lagrange's expansion, but he did not give any related results.

1.3 Organization of this Thesis

This thesis is organized as follows. Chapter 1 provides background information on branching processes. Chapter 2 describes formal power series and coefficient extraction. Chapter 3 describes the Lagrange Inversion Theorem and its application to branching processes. In Chapter 4 and 5, we give the determinant of one special matrix and the distribution of the total progeny for multitype branching process with same offspring distribution. We discuss total progeny for multitype branching process with independent but not identical offspring distribution (Poisson). In chapter 6 we give an application of Dwass' formula.

Chapter 2

The Method of Coefficients

The aim of this chapter is to present the method of extracting the coefficient of formal power series. These techniques will be used in the proof of our main results (Theorem 4.1 and 4.5).

2.1 Formal Power Series

In this section, we will discuss how to construct the ring of formal power series in n complex variables from the ring of formal power series in one complex variable, so that the technique of coefficient extraction for the ring of formal power series in one complex variable can be extended to the case of n complex variables. For the presentation and examples below, we have benefited from the book of Shafarevich[1997].

Definition 2.1. Let $\mathbb{K}$ be a (non-empty) set with two operations, addition (denoted by $a * b$) and multiplication (denoted by $a \cdot b$). $\mathbb{K}$ is said to be a field if the operation satisfy the following conditions for any a, b, c in $\mathbb{K}$:

- Addition
  - Commutativity: $a * b = b * a$
  - Associativity: $a * (b * c) = (a * b) * c$
  - Existence of zero: there exists an element $0 \in \mathbb{K}$ with $a * 0 = a$ for every a (it can be shown that this element is unique)
  - Existence of negative: there exists an element $-a$ with $a * (-a) = 0$ for any a (it can be shown that this element is unique)
- Multiplication
  - Commutativity: $a \cdot b = b \cdot a$
  - Associativity: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
  - Existence of unity: there exists an element $1 \in \mathbb{K}$ with $a \cdot 1 = a$ for every a (it can be shown that this element is unique)
  - Existence of inverse: there exists an element a^{-1} with $a \cdot a^{-1} = 1$ for any $a \neq 0$ (it can be shown that for given a, this element is unique)
- Addition and Multiplication
  - Distributivity: $a \cdot (b * c) = a \cdot b * a \cdot c$

If the requirements of both commutativity and existence of inverse are removed, then $\mathbb{K}$ is said to be a ring. If only the requirements of existence of inverse is removed, $\mathbb{K}$ is said to be a commutative ring.

The above conditions are generally called the field axioms. They will be referred to as the commutative ring axioms if the existence of an inverse and the condition $0 \neq 1$ omitted.

Example 2.1. *(Shafarevich[1997], P.16) Consider the set of all Laurent series* $\sum_{n=-k}^{\infty} a_n z^n$ *which are convergent in an annulus* $0 < |z| < R$ *(where different series may have different annuli of convergence). With the usual definition of operations on series, these form a field, the field of Laurent series. If we use the same rules to compute the coefficients, we can define the sum and product of two Laurent series, even if these are nowhere convergent. We thus obtain the field of formal Laurent series. If the coefficients* a_n *belong to an arbitrary field* $\mathbb{K}$*, the resulting field is called the field of formal Laurent series with coefficients in* $\mathbb{K}$*, and is denoted by* $\mathbb{K}((z))$*,*

Definition 2.2. Let $\mathbb{A}$ be a commutative ring. The set of formal symbols

$$\mathbb{A}[z] = \left\{a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 | a_i \in \mathbb{A}, n \text{ is nonnegative integer}\right\}$$

is called the ring of polynomials over $\mathbb{A}$ in the indeterminate z.

If we denote the sequence $(a_0, a_1, \cdots, a_n)$ as a polynomial $a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$, then sum and product are given by formulas

$$\sum_k a_k z^k + \sum_k b_k z^k = \sum_k (a_k + b_k) z^k \tag{2.1}$$

$$\left(\sum_k a_k z^k\right) \cdot \left(\sum_l b_l z^l\right) = \sum_m c_m z^m \quad \text{where } c_m = \sum_{k+l=m} a_k b_l. \tag{2.2}$$

Consider any infinite sequence $(a_0, a_1, \cdots, a_n, \cdots)$ of elements of a ring $\mathbb{A}$, which consist of zeros from some term onwards (this term may be different for different sequences), then addition of sequences can be defined as

$$(a_0, a_1, \cdots, a_n, \cdots) + (b_0, b_1, \cdots, b_n, \cdots) = (a_0 + b_0, a_1 + b_1, \cdots, a_n + b_n, \cdots)$$

Clearly, all the ring axioms concerning addition are satisfied. As for multiplication, we define first the multiplication of sequences by elements of $\mathbb{A}$:

$$a(a_0, a_1, \cdots, a_n, \cdots) = (aa_0, aa_1, \cdots, aa_n, \cdots)$$

Let $E_k = (0, \cdots, 1, 0, \cdots)$ denote the sequence consisting of 1 in the k^{th} place and 0 everywhere else. Then,

$$(a_0, a_1, \cdots, a_n, \cdots) = \sum_k a_k E_k \tag{2.3}$$

Now define multiplication as follows

$$\left(\sum_k a_k E_k\right) \cdot \left(\sum_l b_l E_l\right) = \sum_{k+l=n} a_k b_l E_{k+l} \tag{2.4}$$

It follows from (2.4) that E_0 is the unit element of the ring, and $E_k = E_1^k$. A unity in a ring is nonzero element u such that there exists a multiplication inverse u^{-1} where $u \cdot u^{-1} = 1$.

Setting $E_1 = z$, the sequence (2.3) can be written in the form $\sum a_k z^k$. Obviously this expression for the sequence is unique. It is easy to check that the multiplication (2.4) satisfies the axioms of a commutative ring, so that the ring we have constructed is the polynomial ring $\mathbb{A}[x]$.

Remark 2.1. The polynomial ring $\mathbb{A}[z_1, z_2]$ is defined as $\mathbb{A}[z_1][z_2]$, or by generalising the above construction. In a similar way one defines the polynomial ring $\mathbb{A}[z_1, \cdots, z_n]$ in any number of variables.

Definition 2.3. An isomorphism of two fields $\mathbb{K}'$ and $\mathbb{K}''$ is a 1-to-1 correspondence $a' \leftrightarrow a''$ between their elements such that $a' \leftrightarrow a''$ and $b' \leftrightarrow b''$ implies that $a' * b' \leftrightarrow a'' * b''$ and $a' \cdot b' \leftrightarrow a'' \cdot b''$; we say that two fields are isomorphic if there exists an isomorphism between them. An isomorphism of fields $\mathbb{K}'$ and $\mathbb{K}''$ is denoted by $\mathbb{K}' \cong \mathbb{K}''$.

Example 2.2. *(Shafarevich[1997], P.19) All linear differential operators with constant (real) coefficients can be written as polynomials in the operators* $\frac{\partial}{\partial z_1}, \cdots, \frac{\partial}{\partial z_1}$. *Hence they form a ring*

$$\mathbb{R}\left[\frac{\partial}{\partial z_1}, \cdots, \frac{\partial}{\partial z_1}\right]$$

Sending $\frac{\partial}{\partial z_i}$ *to* t_i *defines an isomorphism*

$$\mathbb{R}\left[\frac{\partial}{\partial z_1}, \cdots, \frac{\partial}{\partial z_1}\right] \cong \mathbb{R}\left[t_1, \cdots, t_n\right]$$

Definition 2.4. A subset $\mathbb{S}$ of a ring $\mathbb{K}$ is a subring of $\mathbb{K}$ if $\mathbb{S}$ is itself a ring with operations of $\mathbb{K}$.

Definition 2.5. A commutative ring with the properties that for any elements a, b the product $a \cdot b = 0$ only if $a = 0$ or $b = 0$, and that $0 \neq 1$, is called an integral ring or an integral domain. Thus a subring of any field is an integral domain.

Theorem 2.1. *For any integral domain* $\mathbb{A}$, *there exists a field* $\mathbb{K}$ *containing* $\mathbb{A}$ *as a subring, and such that every element of* $\mathbb{K}$ *can be written in the form* $a \cdot b^{-1}$ *with* $a, b \in \mathbb{A}$ *and* $b \neq 0$. *A field* $\mathbb{K}$ *with this property is called the field of fraction of* $\mathbb{A}$: *it is uniquely defined up to isomorphism.*

Remark 2.2. All rational functions form a field, called the rational function field; it is denoted by $\mathbb{K}(z)$.

Remark 2.3. The field of fractions of the polynomial ring $\mathbb{K}[z]$ is the field of rational functions $\mathbb{K}(z)$, and that of $\mathbb{K}[z_1, \cdots, z_n]$ is $\mathbb{K}(z_1, \cdots, z_n)$

Definition 2.6. Let $\mathbb{R}$ be the field of real numbers and let z be any indeterminate over $\mathbb{R}$, i.e., a symbol different from any element in $\mathbb{R}$. A formal power series over $\mathbb{R}$ in the indeterminate z is an expression:

$$F(z) = \sum_{n\geq 0} f_n z^n$$

for real-valued coefficients f_n. In this case, f_n is called the n^{th} coefficient of $F(z)$, write $[z^n]F = f_n$. If $f_0 = 0$, $F(z)$ is called a nonunit.

Two formal power series can be added by adding the coefficients of like powers:

$$\sum_{n\geq 0} a_n z^n + \sum_{n\geq 0} b_n z^n = \sum_{n\geq 0} (a_n + b_n) z^n \tag{2.5}$$

Two formal power series can be multiplied:

$$\left(\sum_{n\geq 0} a_n z^n\right) \cdot \left(\sum_{n\geq 0} b_n z^n\right) = \sum_{n\geq 0} c_n z^n \tag{2.6}$$

Where $c_n = \sum_{j=0}^{n} a_j b_{n-j}$.

Example 2.3. *(Shafarevich[1997], p.20) The ring of functions of one complex variable holomorphic at the origin is an integral domain, and its field of fractions is the field of Laurent series. Similarly to Example 2.1 we can define the ring of formal power series* $\sum_{n=0}^{\infty} a_n z^n$ *with coefficients* a_n *in any field* $\mathbb{K}$. *This can also be constructed as in Example 2.1, if we just omit the condition that the sequences* $(a_0, a_1, \cdots, a_n, \cdots)$ *are 0 from some point onwards. This is also an integral domain, and its field of fractions is the field of formal Laurent series* $\mathbb{K}((z))$. *The ring of formal power series is denoted by* $\mathbb{K}[[z]]$.

Definition 2.7. Let $\mathcal{F}$ be a ring, let $d \in \mathbb{N}$ be given, a formal power series on $\mathcal{F}$ is defined to be a map $F : \mathbb{N}^d \to \mathcal{F}$

$$F := \sum_{\underline{n}} f_{\underline{n}} \underline{z}^{\underline{n}}$$

where $f_{\underline{n}}$ denotes $f_{i_1 \cdots i_d}$ and $\underline{z}^{\underline{n}}$ denotes $z_1^{n_1} \cdots z_d^{n_d}$.

As above, We will write $[\underline{z}^{\underline{n}}]F$ to refer to the coefficient of $\underline{z}^{\underline{n}}$ in F.

Given two formal power series $A(\underline{z})$ and $B(\underline{z})$ we will define their sum and product respectively as

$$A + B := \sum_{\underline{n}} u_{\underline{n}} \underline{z}^{\underline{n}} \tag{2.7}$$

$$AB := \sum_{\underline{n}} v_{\underline{n}} \underline{z}^{\underline{n}} \tag{2.8}$$

where

$$u_{\underline{n}} := a_{\underline{n}} + b_{\underline{n}} \quad \textit{and} \quad \underline{v}_{\underline{n}} := \sum_{\underline{k}:\ \underline{0 \le k \le n}} a_{\underline{k}} b_{\underline{n}-\underline{k}}$$

Example 2.4. *(Shafarevich[1997], P.21) The ring $\mathcal{O}_n$ of functions in n complex variables holomorphic at the origin, that is of functions that can be represented as power series*

$$\sum a_{i_1 \cdots i_n} z_1^{i_1} \cdots z_n^{i_n}$$

convergent in some neighbourhood of the origin. By analogy with Example 2.3 we can define the rings of formal power series $\mathbb{C}[[z_1, \cdots, z_n]]$ with complex coefficients, and $\mathbb{K}[[z_1, \cdots, z_n]]$ with coefficients in any field $\mathbb{K}$

2.2 Composition

Consider two formal power series $G(z) = b_0 + b_1 z + b_2 z^2 + \cdots$ and $F(z) = a_1 z + a_2 z^2 + \cdots$, it is well known that the composition of G with F , $G(F(z))$, is again a formal power series. If the constant term of F is not zero, then the composition $G(F(z))$ may not exist.

Definition 2.8. Let $F(z) = \sum_{n=0}^{\infty} f_n z^n$ be a formal power series. The order of F is the least integer n for which $f_n \neq 0$, and is denoted by ord(F). The norm,$\| \cdot \|$, of F is defined as $\| F \| = 2^{ord(F)}$, except that the norm of the zero formal power series is defined to be zero.

Definition 2.9. Let $\mathcal{F}$ be a ring and let $\mathbb{F}$ be the set of all formal power series over $\mathcal{F}$. Let $G(z) \in \mathbb{F}$ be given, say $G(z) = \sum_{n=0}^{\infty} g_n z^n$. We define a subset $\mathbb{F}_G \subset \mathbb{F}$ to be

$$\mathbb{F}_G = \left\{ F(z) \in \mathbb{F} \mid F(z) = \sum_{n=0}^{\infty} f_n z^n, \sum_{k=0}^{\infty} g_k f_n^{(k)} \in \mathcal{F} \quad n = 0, 1, 2, \cdots \right\}$$

where $F^k(z) = \sum_{n=0}^{\infty} f_n^{(k)} z^n$ for all $k \in \mathbb{N}$, created by the product rule in Definition 2.6. We will see that $\mathbb{F}_G \neq \emptyset$ by Proposition 2.2. Then the mapping $T_G : \mathbb{F}_G \to \mathbb{F}$ such that

$$T_G(F)(z) = \sum_{n=0}^{\infty} c_n z^n$$

where $c_n = \sum_{k=0}^{\infty} g_k f_n^{(k)}, n = 0, 1, 2, \cdots$, is well defined. We call $T_G(F)$ the composition of G and F ; $T_G(F)$ is also denoted by $G \circ F$.

Example 2.5. *(Gan and Knox[2002]) Let $\mathcal{F} = \mathbb{R}$. Let $G(z) = \sum_{n=0}^{\infty} z^n$ and $F(z) = 1+z$. We cannot calculate even the first coefficient of the series $\sum_{n=0}^{\infty} (F(z))^n$ under Definition 2.8. Thus, the composition G(F(z)) does not exist.*

Under these definitions, a composition was established as follows.

Proposition 2.2. *(Roman[1992]) Let* $F(z) = \sum_{n=0}^{\infty} f_n z^n$ *be a formal power series in z. If G is a formal power series, such that,*

$$\lim_{n \longrightarrow \infty} \|f_n G^n\| = 0$$

then the $\sum f_n G^n$ *converges to a power series. This series is called the composition of F and G and is denoted by* $F \circ G$.

Clearly, the requirement $\lim_{n \longrightarrow \infty} \|f_n G^n\| = 0$ implies that the only candidates for such G are formal power series such that $G(0) = 0$ unless F is a polynomial. The most recent progress on the existence of the composition of formal power series can be found in Chaumat and Chollet [2001] where they discussed the radius of convergence of composed formal power series and obtained some very good results.

In our case, we have

$$\begin{aligned} w^{(i)}(\underline{0}) &= s_i f^{(i)}\left(w^{(1)}(\underline{s}), \cdots, w^{(d)}(\underline{s})\right) \qquad i = 1, \cdots, d \\ &= 0 \end{aligned}$$

2.3 Coefficient Extraction

A formal power series can be seen as a sequence of its coefficients. The composition of formal power series is eventually determined by its coefficients.

Given a formal power series $F(z) = \sum_{n=0}^{\infty} f_n z^n$, it is often important to be able to determine the coefficient of z^n in $F(z)$ for some n. The notation $[z^n]F(z)$ indicates the extraction of the coefficient of z^n from $F(z)$, therefore, $[z^n]F(z) = f_n$.

The following results are well-known and appear in various papers. We don't know who first gave proofs, but it is easy to prove them.

Theorem 2.3. *Let* $F(z) = \sum_n f_n z^n$ *and* $G(z) = \sum_n g_n z^n$ *be formal power series. Then*

1. *Linearity:* $[z^n]\left(F(z) + G(z)\right) = ([z^n]F(z)) + ([z^n]G(z))$.

2. *Let* $c \in \mathbb{R}$, *then* $[z^n]\left(cF(z)\right) = c[z^n]F(z)$.

3. *Scaling: If* c *is a constant,* $[z^n]F(cz) = c^n[z^n]F(z)$

4. *Right-shifting:* $[z^n]z^k F(z) = [z^{n-k}]F(z)$

5. *Left-shifting*

   For one-sided series, we can create a new series by

   $$H(z) = \frac{F(z) - \sum_{n=0}^{m-1} f_n z^n}{z^m}$$

   That is, $[z^n]H(z) = [z^{n+m}]F(z) for\ n \geq 0$. *This is a truncated left shift, and the sum above cannot in general be extended over all integers. A two-sided left-shift is obtained by* $[z^n]F(z)/z^m = [z^{n+m}]F(z)$; *this is valid for all* m *but usually less useful.*

6. *Differentiation: Let* $F'(z) = \sum n f_n z^{n-1}$, *then* $[z^{n-1}]F'(z) = n[z^n]F(z)$

We define the partial derivative of $F(\underline{z})$ with respect to z_j as the formal power series defined as

$$\frac{\partial F}{\partial z_j} = \sum_{\underline{n}:n_j \geq 1} n_j f_{\underline{n}} \underline{z}^{\underline{n}-\underline{e}_j} \tag{2.9}$$

where $\underline{e}_j$ is the vector that has all its coordinates identically zero, however, its j^{th} coordinate equals 1.

Corollary 2.4.

$$[\underline{z}^{\underline{n}-\underline{e}_j}]\frac{\partial F}{\partial z_j} = n_j[\underline{z}^{\underline{n}}]F$$

If we regard $\mathbb{C}[[\underline{z}]]$ as a vector space over $\mathbb{C}$ then $\frac{\partial}{\partial z_j}$ is a linear operator. Since the operators $\frac{\partial}{\partial z_j}$ and $\frac{\partial}{\partial z_k}$ commute we can unambiguously define, for all $\underline{k}$, the pseudo-derivative

$$\frac{\partial^{\underline{k}} F}{\partial \underline{z}_{\underline{k}}} = \frac{\partial^{k_1}}{\partial z_1^{k_1}} \cdots \frac{\partial^{k_d}}{\partial z_1^{k_d}} F = \sum_{\underline{n}:\underline{n} \geq \underline{k}} \frac{\underline{n}!}{(\underline{n}-\underline{k})!} f_{\underline{n}} \underline{z}^{\underline{n}-\underline{k}} \tag{2.10}$$

In particular, for all $\underline{k}$, we have the identity

Corollary 2.5.

$$[\underline{z}^{\underline{k}}]F = \frac{1}{\underline{k}!}\left\{\frac{\partial^{\underline{k}} F}{\partial \underline{z}_{\underline{k}}}\right\}_{\underline{z}=\underline{0}} \tag{2.11}$$

Chapter 3

Lagrange Inversion Theorem with Application to Branching Processes

Given a formal power series, the determination of its compositional inverse is one of the most interesting problems; it was solved by Lagrange and we will discuss it in the following sections.

3.1 Lagrange Inversion Formula

Definition 3.1. A function $f(z)$ of one complex variable is analytic in a connected open set $A \subset \mathbb{C}$ if in a small neighborhood of every point $\omega \in A$, $f(z)$ has an expansion as a power series

$$f(z) = \sum_{n=0}^{\infty} a_n (z-\omega)^n, \quad a_n = a_n(\omega) \tag{3.1}$$

that converges.

Definition 3.2. A function $f(z)$ is called meromorphic in A if it is analytic in A except at a (countable isolated) subset $A' \subset A$, and in a small neighborhood of every $\omega \in A'$, $f(z)$ has an expansion of the form

$$f(z) = \sum_{n=-N(\omega)}^{\infty} a_n (z-\omega)^n, \quad a_n = a_n(\omega) \tag{3.2}$$

Thus meromorphic functions can have poles.

Theorem 3.1. *Lagrange's Theorem*

Let $f(z)$ and $\phi(z)$ be function of z analytic on and inside a contour C surrounding a point a, and let t be such that the inequality

$$|t\phi(z)| < |z-a|$$

is satisfied at all points on the perimeter of C; then the equation

$$\zeta = a + t\phi(\zeta)$$

regard as an equation in ζ, has one root in the interior of C; and further any function of ζ analytic on and inside C can be expanded as a power series in t by the formula

$$f(\zeta) = f(a) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \frac{d^{n-1}}{dz^{n-1}} \{f'(z)\,[\phi(z)]^n\}_{z=a} \tag{3.3}$$

Theorem 3.2. *Lagrange-Bürmann Inversion Formula*

Let f(z) be defined implicitly by the equation $f(z) = z\varphi(f(z))$, where $\varphi(u)$ is a series with $\varphi(0) \neq 0$. Then the coefficients of f(z), its powers $f^k(z)$, and an arbitrary composition $g(f(z))$ are related to the coefficients of the powers of $\varphi(u)$ as follows:

$$[z^n]f(z) = \frac{1}{n}[u^{n-1}]\varphi^n(u) \tag{3.4}$$

$$[z^n]f^k(z) = \frac{k}{n}[u^{n-k}]\varphi^n(u) \tag{3.5}$$

$$[z^n]g\left(f(z)\right) = \frac{1}{n}[u^{n-1}]\varphi^n(u)g'(u) \tag{3.6}$$

Many different versions of multivariate Lagrange inversion formulas have been found, such as those of Jacobi, Stieltjes, Good, Joni and Abhyankar (Gessel [1987]). The following two theorems are due to Good[1960].

Theorem 3.3. *If $h(\underline{z})$ is analytic in a neighborhood of $\underline{z} = \underline{a}$, if*

$$\zeta_\mu - \alpha_\mu = \frac{z_\mu - a_\mu}{f_\mu\left(\underline{z}\right)} \quad \mu = 1, 2, \cdots, d$$

where $f_\mu\left(\underline{a}\right) \neq 0$, then

$$h\left(\underline{z}\left(\underline{\zeta}\right)\right) = \sum \frac{(\zeta_1 - \alpha_1)^{m_1} \cdots (\zeta_d - \alpha_d)^{m_d}}{m_1! \cdots m_d!} \left\{ \frac{\partial^{m_1 + \cdots + m_d}}{\partial t_1^{m1} \cdots \partial t_d^{m_d}} \left[H(\underline{t}) \left(f_1(\underline{t})\right)^{m_1} \cdots \left(f_d(\underline{t})\right)^{m_d}\right] \right\}_{\underline{t}=\underline{a}}$$

where

$$H(\underline{t}) = h(\underline{t}) \left\| \delta_\mu^\nu - \frac{t_\mu - a_\mu}{f_\mu\left(\underline{t}\right)} \frac{\partial f_\mu\left(\underline{t}\right)}{\partial t_\nu} \right\|$$

Theorem 3.4. *Multivariable Lagrange Inversion Formula*

If $\zeta_\mu = \frac{z_\mu}{f_\mu(\underline{z})}$, where $f_\mu\left(\underline{z}\right)$ is analytic in a neighbourhood of the origin and $f_\mu\left(\underline{0}\right) \neq 0$ for $\mu = 1, \cdots, d$), and $h(\underline{z})$ is meromorphic in a neighbourhood of the origin, then

$$[\zeta^{m_1} \cdots \zeta^{m_d}]h\left(\underline{z}\left(\underline{\zeta}\right)\right) = [z^{m_1} \cdots z^{m_d}]h(\underline{z}) \left(f_1(\underline{z})\right)^{m_1} \cdots \left(f_d(\underline{z})\right)^{m_d} \left\| \delta_\mu^\nu - \frac{z_\mu}{f_\mu\left(\underline{z}\right)} \frac{\partial f_\mu\left(\underline{z}\right)}{\partial z_\nu} \right\|$$

where δ_μ^ν is the Kronecker delta, and $\|a_{jk}\|$ denotes the determinant of the matrix (a_{jk}).

We note that for $d = 1$

$$[\zeta^m]h\left(z\left(\zeta\right)\right) = [z^m]h(z) \left(f^m(z) \left(1 - z\frac{f'\left(z\right)}{f\left(z\right)}\right)\right) \tag{3.7}$$

3.2 Application

When Theorems 2.3 and 3.2 are applied to a single type branching process, it is easy to obtain Dwass' formula. Let the branching process start with k individuals in the zeroth generation. Let $g(s) = s^i$ and $w = sf(w)$. By Lemma 1.2, $g(w)$ is the generating function for the total numbers of the various types in all generations.

$$
\begin{aligned}
[s^m]g(w(s)) &= \frac{1}{m}[u^{m-1}]f^n(u)g'(u) \\
&= \frac{i}{m}[u^{m-1}]u^{i-1}f^m(u) \\
&= \frac{i}{m}[u^{m-i}]f^m(u)
\end{aligned}
$$

Good[1960] gave the following result. Let the branching process start with $i_1, \cdots, i_d$ individuals in the zeroth generation, of types 1, 2, $\cdots$, d. Let $h(\underline{s}) = s_1^{i_1} \cdots s_d^{i_d}$ and $w = \underline{s}$, by Lemma 1.2, $h(w)$ is the generating function for the total numbers of the various types in all generations. Apply theorem 2.2 and 4.4, then obtain

Proposition 3.5. *(Good[1960]) If the branching process starts off with i_1 individuals of type 1, i_2 of type 2, $\cdots$, i_d of type d, then the probability that the whole process will contain precisely m_1 individuals of type 1, $\cdots$, m_d of type d, is equal to the coefficients of $s^{m_1-i_1} \cdots s^{m_d-i_d}$ in*

$$
\left(f^{(1)}(\underline{s})\right)^{m_1} \cdots \left(f^{(d)}(\underline{s})\right)^{m_d} \left\| \delta_\mu^\nu - \frac{s_\mu}{f^{(\mu)}(\underline{s})} \frac{\partial f^{(\mu)}(\underline{s})}{\partial s_\nu} \right\|
$$

That is,

$$
P(\underline{S} = \underline{m} \mid \underline{Z}_0 = \underline{i}) = [s^{m_1-i_1} \cdots s^{m_d-i_d}] \left(f^{(1)}(\underline{s})\right)^{m_1} \cdots \left(f^{(d)}(\underline{s})\right)^{m_d} \left\| \delta_\mu^\nu - \frac{s_\mu}{f^{(\mu)}(\underline{s})} \frac{\partial f^{(\mu)}(\underline{s})}{\partial s_\nu} \right\| \tag{3.8}
$$

The conditions $f^{(\mu)}(\underline{0}) \neq 0$ $(\mu = 1, \cdots, d)$ are required for Theorem 3.4. Good also pointed out that if the branching process is finite we must not have $f^{(\mu)}(\underline{0}) = 0$ for all μ.

In fact, Good almost found the same result as Dwass. When $d = 1$, we have

$$P(S = m \mid Z_0 = i) = [s^{m-i}]f^m(s)\left(1 - s\frac{f'(s)}{f(s)}\right)$$

Note that

$$\begin{aligned} f^m(s)\left(1 - s\frac{f'(s)}{f(s)}\right) &= f^m(s) - sf^{m-1}(s)f'(s) \\ &= f^m(s) - \frac{s}{m}\left(f^m(s)\right)' \end{aligned}$$

By Theorem 2.3 , we have

$$\begin{aligned} [s^{m-i}]\frac{s}{m}\left(f^m(s)\right)' &= \frac{1}{m}[s^{m-i-1}]\left(f^m(s)\right)' \\ &= \frac{1}{m}[s^{m-i}](m-i)\left(f^m(s)\right) \end{aligned}$$

Therefore, we obtain

$$P(S = m \mid Z_0 = i) = [s^{m-i}]f^m(s) - \frac{1}{m}[s^{m-i}](m-i)\left(f^m(s)\right) = \frac{i}{m}[s^{m-i}]f^m(s)$$

Chapter 4

Main Results

It is of course natural to ask whether a multitype case has a similar formula to Dwass. When each type has identical offspring distributions, the following theorem extends Dwass' formula to multitype branching process, but I can not yet extend it to case of different offspring distribution in this thesis.

Theorem 4.1. *If the branching process starts off with i_1 individuals of type 1, i_2 of type 2, $\cdots$, i_d of type d, and suppose that the distributions of children are in the sense that*

$$f^{(\kappa)}(s_1, \cdots, s_d) = F(s_1, \cdots, s_d) \qquad \kappa = 1, 2, \cdots, d$$

with $f^{(\kappa)}(0, \cdots, 0) \neq 0$. Then the probability that the whole process will contain precisely m_1 individuals of type 1, $\cdots$, m_d of type d, is given by

$$P(\underline{S} = \underline{m} \mid \underline{Z}_0 = \underline{i}) = \frac{i_1 + \cdots + i_d}{m_1 + \cdots + m_d} P\left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d)\right)$$

From the above theorem, we immediately have

$$\omega^{(j)}(s_1, \cdots, s_d) \\ = \frac{1}{m_1 + \cdots + m_d} P\left(\underline{Z}_1 = (m_1, \cdots, m_{j-1}, m_j - 1, m_{j+1}, \cdots, m_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d)\right) \tag{4.1}$$

To prove Theorem 4.1, we need the following result.

Theorem 4.2. *Let* $A = (a_1, \cdots, a_n)^T$ *and* $S = (s_1, \cdots, s_n)$*, Define*

$$M^{(n)} = I - AS = \begin{pmatrix} 1 - a_1 s_1 & -a_2 s_1 & \cdots & -a_n s_1 \\ -a_1 s_2 & 1 - a_2 s_2 & \cdots & -a_n s_2 \\ \cdots & \cdots & \cdots & \cdots \\ -a_1 s_n & -a_2 s_n & \cdots & 1 - a_n s_n \end{pmatrix}$$

then

$$M_{11}^{(n)} = 1 - \sum_{l=2}^{n} a_l s_l \tag{4.2}$$

$$M_{j1}^{(n)} = (-1)^{j-1} a_j s_1 \qquad j = 2, 3, \cdots, n \tag{4.3}$$

and so

$$det(I - AS) = \|M^{(n)}\| = 1 - A \bullet S = 1 - \sum_{l=1}^{n} a_l s_l \tag{4.4}$$

First, we review some definitions and properties of determinants, which are found in most linear algebra textbooks. They will be needed for the proof of Theorem 4.2.

Denote by $M_{m\times n}$ the set of $m \times n$ complex matrices and by $M^{(n)}$ the set $M_{n\times n}$. Before discussing the computation of determinants using cofactors a few definitions concerning matrices and submatrices will be useful.

Definition 4.1. A submatrix is a matrix formed from a matrix M by taking a subset consisting of j rows with column elements from a set of k columns.

Definition 4.2. A minor is the determinant of a square submatrix of the matrix M.

Definition 4.3. The minor associated with the element m_{ij} of a square matrix $M^{(n)}$, denoted by $M_{ij}^{(n)}$, is obtained by including all but the i^{th} row and the j^{th} column, or alternatively the minor that is obtained by deleting the i^{th} row and the j^{th} column.

Lemma 4.3. *The determinant of a square matrix $M^{(n)} = (m_{ij}^{(n)})$ can be defined over $M^{(n-1)}$,*

$$\|M^{(n)}\| = \sum_{j=1}^{n} (-1)^{i+j} m_{ij}^{(n)} M_{ij}^{(n)}$$

where i denotes the i^{th} row of the matrix M. This is called an expansion of $\|M\|$ by column i of M. The result is the same for any other row. This can also be done for columns letting the sum range over i instead of j.

Lemma 4.4. *The sign of the determinant will change if we interchange two rows (or two columns).*

Proof of Theorem 4.2. We use induction to prove it.

1. First check $n = 1, 2, 3$

   - $n = 1$, Clearly we have

     $$\|I - AS\| = 1 - as$$

   - $n = 2$

     $$M^{(2)} = I - AS = \begin{pmatrix} 1 - a_1 s_1 & -a_2 s_1 \\ -a_1 s_2 & 1 - a_2 s_2 \end{pmatrix}$$

     Clearly we have

     $$\begin{aligned} M_{11}^{(2)} &= 1 - a_2 s_2 \\ M_{21}^{(2)} &= -a_2 s_1 \end{aligned}$$

and

$$\|M^{(2)}\| = 1 - a_1 s_1 - a_2 s_2$$

- $n = 3$

$$M^{(3)} = I - AS = \begin{pmatrix} 1 - a_1 s_1 & -a_2 s_1 & -a_3 s_1 \\ -a_1 s_2 & 1 - a_2 s_2 & -a_3 s_2 \\ -a_1 s_3 & -a_2 s_3 & 1 - a_3 s_3 \end{pmatrix}$$

Clearly we have

$$M_{11}^{(3)} = \left\| \begin{matrix} 1 - a_2 s_2 & -a_3 s_2 \\ -a_2 s_3 & 1 - a_3 s_3 \end{matrix} \right\| = 1 - a_2 s_2 - a_3 s_3$$

$$M_{21}^{(3)} = \left\| \begin{matrix} -a_2 s_1 & -a_3 s_1 \\ -a_2 s_3 & 1 - a_3 s_3 \end{matrix} \right\| = -a_2 s_1 (1 - a_3 s_3) - a_2 a_3 s_1 s_3 = -a_2 s_1$$

$$M_{31}^{(3)} = \left\| \begin{matrix} -a_2 s_1 & -a_3 s_1 \\ 1 - a_2 s_2 & -a_3 s_2 \end{matrix} \right\| = a_2 a_3 s_1 s_2 + a_3 s_1 (1 - a_2 s_2) = a_3 s_1$$

By Lemma 4.3, we have

$$\begin{aligned} \|M^{(3)}\| &= (-1)^{1+1}(1 - a_1 s_1) M_{11} + (-1)^{1+2}(-a_1 s_2) M_{21} + (-1)^{1+3}(-a_1 s_3) M_{31} \\ &= (1 - a_1 s_1)(1 - a_2 s_2 - a_3 s_3) - a_1 a_2 s_1 s_2 - a_1 a_3 s_1 s_3 \\ &= 1 - a_1 s_1 - a_2 s_2 - a_3 s_3 \end{aligned}$$

2. Next, suppose that (4.2),(4.3) and (4.4) are true for $n = k$. We need to show

that they are also true for $n = k + 1$. We write

$$M^{(k+1)} = \begin{pmatrix} 1 - a_1 s_1 & -a_2 s_1 & \cdots & -a_{k+1} s_1 \\ -a_1 s_2 & 1 - a_2 s_2 & \cdots & -a_{k+1} s_2 \\ \cdots & \cdots & \cdots & \cdots \\ -a_1 s_{k+1} & -a_2 s_{k+1} & \cdots & 1 - a_{k+1} s_{k+1} \end{pmatrix}$$

Again, by Lemma 4.3, we have

$$\|M^{(k+1)}\| = (1 - a_1 s_1) M_{11}^{(k+1)} + \sum_{j=2}^{k+1} (-1)^{1+j} (-a_1 s_j) M_{j1}^{(k+1)}$$

where

$$M_{11}^{(k+1)} = \begin{Vmatrix} 1 - a_2 s_2 & -a_3 s_2 & \cdots & -a_{k+1} s_2 \\ -a_2 s_3 & 1 - a_3 s_3 & \cdots & -a_{k+1} s_3 \\ \cdots & \cdots & \cdots & \cdots \\ -a_2 s_{k+1} & -a_3 s_{k+1} & \cdots & 1 - a_{k+1} s_{k+1} \end{Vmatrix}$$

$$M_{21}^{(k+1)} = \begin{Vmatrix} -a_2 s_1 & -a_3 s_1 & \cdots & -a_{k+1} s_1 \\ -a_2 s_3 & 1 - a_3 s_3 & \cdots & -a_{k+1} s_3 \\ \cdots & \cdots & \cdots & \cdots \\ -a_2 s_{k+1} & -a_3 s_{k+1} & \cdots & 1 - a_{k+1} s_{k+1} \end{Vmatrix}$$

For $3 \leq j \leq k$

$$M_{j1}^{(k+1)} = \begin{Vmatrix} -a_2s_1 & -a_3s_1 & \cdots & \cdots & \cdots & -a_{k+1}s_1 \\ 1-a_2s_2 & -a_3s_2 & \cdots & \cdots & \cdots & -a_{k+1}s_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_2s_{j-1} & \cdots & 1-a_{j-1}s_{j-1} & \cdots & \cdots & -a_{k+1}s_{j-1} \\ -a_2s_{j+1} & \cdots & \cdots & -a_js_{j+1} & \cdots & -a_{k+1}s_{j+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_2s_{k+1} & \cdots & \cdots & \cdots & \cdots & 1-a_{k+1}s_{k+1} \end{Vmatrix}$$

and $j = k + 1$

$$M_{(k+1)1}^{(k+1)} = \begin{Vmatrix} -a_2s_1 & -a_3s_1 & \cdots & \cdots & -a_{k+1}s_1 \\ 1-a_2s_2 & -a_3s_2 & \cdots & \cdots & -a_{k+1}s_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_2s_k & -a_3s_k & \cdots & 1-a_ks_k & -a_{k+1}s_k \end{Vmatrix}$$

Since

$$M_{21}^{(k+1)} = \left\|\begin{matrix} -a_2 s_1 & -a_3 s_1 & \cdots & a_{k+1} s_1 \\ -a_2 s_3 & 1 - a_3 s_3 & \cdots & -a_{k+1} s_3 \\ \cdots & \cdots & \cdots & \cdots \\ -a_2 s_{k+1} & -a_3 s_{k+1} & \cdots & 1 - a_{k+1} s_{k+1} \end{matrix}\right\|$$

$$= -a_2 s_1 \left\|\begin{matrix} 1 - a_3 s_3 & -a_4 s_3 & \cdots & -a_{k+1} s_3 \\ -a_3 s_4 & 1 - a_4 s_4 & \cdots & -a_{k+1} s_4 \\ \cdots & \cdots & \cdots & \cdots \\ -a_3 s_{k+1} & -a_4 s_{k+1} & \cdots & 1 - a_{k+1} s_{k+1} \end{matrix}\right\|$$

$$+ \sum_{l=3}^{k} (-1)^{l-1} a_2 s_l \left\|\begin{matrix} -a_3 s_1 & -a_4 s_1 & \cdots & \cdots & \cdots & \cdots & a_{k+1} s_1 \\ 1 - a_3 s_3 & -a_4 s_3 & \cdots & \cdots & \cdots & \cdots & -a_{k+1} s_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ -a_3 s_{l-1} & -a_4 s_{l-1} & \cdots & 1 - a_{l-1} s_{l-1} & \cdots & \cdots & -a_{k+1} s_{l-1} \\ -a_3 s_{l+1} & -a_4 s_{l+1} & \cdots & \cdots & -a_l s_{l+1} & \cdots & -a_{k+1} s_{l+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_3 s_{k+1} & -a_4 s_{k+1} & \cdots & \cdots & \cdots & \cdots & 1 - a_{k+1} s_{k+1} \end{matrix}\right\|$$

$$+ (-1)^k a_2 s_{k+1} \left\|\begin{matrix} -a_3 s_1 & -a_4 s_1 & \cdots & \cdots & a_{k+1} s_1 \\ 1 - a_3 s_3 & -a_4 s_3 & \cdots & \cdots & -a_{k+1} s_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_3 s_k & -a_3 s_k & \cdots & 1 - a_k s_k & -a_{k+1} s_k \end{matrix}\right\|$$

Put

$$N_{11}=\left\|\begin{matrix} 1-a_3s_3 & -a_4s_3 & \cdots & -a_{k+1}s_3 \\ -a_3s_4 & 1-a_4s_4 & \cdots & -a_{k+1}s_4 \\ \cdots & \cdots & \cdots & \cdots \\ -a_3s_{k+1} & -a_4s_{k+1} & \cdots & 1-a_{k+1}s_{k+1} \end{matrix}\right\|$$

For $l=3,\cdots,k$

$$N_{l1}=\left\|\begin{matrix} -a_3s_1 & -a_4s_1 & \cdots & \cdots & \cdots & \cdots & a_{k+1}s_1 \\ 1-a_3s_3 & -a_4s_3 & \cdots & \cdots & \cdots & \cdots & -a_{k+1}s_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ -a_3s_{l-1} & -a_4s_{l-1} & \cdots & 1-a_{l-1}s_{l-1} & \cdots & \cdots & -a_{k+1}s_{l-1} \\ -a_3s_{l+1} & -a_4s_{l+1} & \cdots & \cdots & -a_ls_{l+1} & \cdots & -a_{k+1}s_{l+1} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_3s_{k+1} & -a_4s_{k+1} & \cdots & \cdots & \cdots & \cdots & 1-a_{k+1}s_{k+1} \end{matrix}\right\|$$

and

$$N_{(k+1)1}=\left\|\begin{matrix} -a_3s_1 & -a_4s_1 & \cdots & \cdots & a_{k+1}s_1 \\ 1-a_3s_3 & -a_4s_3 & \cdots & \cdots & -a_{k+1}s_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_3s_k & -a_3s_k & \cdots & 1-a_ks_k & -a_{k+1}s_k \end{matrix}\right\|$$

Suppose that

$$W^{(k)}=\begin{pmatrix} 1-a_2s_1 & -a_3s_1 & -a_4s_1 & \cdots & a_{k+1}s_1 \\ -a_2s_3 & 1-a_3s_3 & -a_4s_3 & \cdots & -a_{k+1}s_3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -a_2s_{k+1} & -a_3s_{k+1} & -a_4s_{k+1} & \cdots & 1-a_{k+1}s_{k+1} \end{pmatrix}$$

We observe that

$$\begin{aligned} N_{11} &= W_{11}^{(k)} = 1 - \sum_{l=3}^{k+1} a_l s_l \\ N_{l1} &= W_{(l-1)1}^{(k)} = (-1)^{l-2} a_l s_1 \qquad l = 3, \cdots, k+1 \end{aligned}$$

Therefore

$$M_{21}^{(k+1)} = -a_2 s_1 (1 - \sum_{l=3}^{k+1} a_l s_l) + \sum_{l=3}^{k+1} (-1)^{2l-1} a_2 s_1 a_l s_l = -a_2 s_1$$

We now consider $M_{j1}^{(k+1)}$ for $j = 3, 4, \cdots, k+1$. We observe that $M_{j1}^{(k+1)}$ has the form of $M_{21}^{(k+1)}$ after $j-2$ column interchanges. That is, the $(j-1)^{th}$ column, with the entry $-a_j s_{j+1}$, is interchanged with the $(j-2)^{th}$ column. Continue this process until the original $(j-1)^{th}$ column locates to the first column. There are a total of $j-2$ column interchanges. So we get

$$M_{j1}^{(k+1)} = (-1)^{j-2}(-1) a_j s_1 = (-1)^{j-1} a_j s_1 \quad j = 3, \cdots, k+1$$

Also, by our assumption

$$M_{11}^{(k+1)} = 1 - \sum_{l=2}^{k+1} a_l s_l$$

Therefore

$$\begin{aligned} \|M^{(k+1)}\| &= (1 - a_1 s_1)(1 - \sum_{l=2}^{k+1} a_l s_l) + \sum_{j=2}^{k+1} (-1)^{1+j} (-a_1 s_j)(-1)^{j-1} a_j s_1 \\ &= (1 - a_1 s_1)(1 - \sum_{l=2}^{k+1} a_l s_l) - \sum_{j=2}^{k+1} a_1 a_j s_1 s_j \\ &= 1 - \sum_{l=1}^{k+1} a_l s_l + \sum_{l=2}^{k+1} a_1 a_l s_1 s_l - \sum_{j=2}^{k+1} a_1 a_j s_1 s_j \\ &= 1 - \sum_{l=1}^{k+1} a_l s_l \end{aligned}$$

This validates our assumption for $n = k + 1$ and completes the proof.

□

Proof of Theorem 4.1. Note from Theorem 3.4

$$\left\| \delta_\mu^\nu - \frac{s_\mu}{f^{(\mu)}(\underline{s})} \frac{\partial f^{(\mu)}(\underline{s})}{\partial s_\nu} \right\| = \left\| \begin{matrix} 1 - \frac{s_1}{F}\frac{\partial F}{\partial s_1} & -\frac{s_1}{F}\frac{\partial F}{\partial s_2} & \cdots & -\frac{s_1}{F}\frac{\partial F}{\partial s_d} \\ -\frac{s_2}{F}\frac{\partial F}{\partial s_1} & 1 - \frac{s_2}{F}\frac{\partial F}{\partial s_2} & \cdots & -\frac{s_2}{F}\frac{\partial F}{\partial s_d} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{s_d}{F}\frac{\partial F}{\partial s_1} & -\frac{s_d}{F}\frac{\partial F}{\partial s_2} & \cdots & 1 - \frac{s_d}{F}\frac{\partial F}{\partial s_d} \end{matrix} \right\|$$

$$= 1 - \sum_{l=1}^{d} \frac{s_l}{F} \frac{\partial F}{\partial s_l}$$

where the last equality comes from Theorem 4.2. We also know that

$$P\left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d)\right) = \left[s^{m_1 - i_1} \cdots s^{m_d - i_d}\right] \left(F(\underline{s})\right)^{m_1 + \cdots + m_d}$$

On the other hand, by Theorem 2.3 and Corollary 2.4, we have

$$\begin{aligned}
&\left[s^{m_1-i_1} \cdots s^{m_d-i_d}\right] \left(F(\underline{s})\right)^{m_1+\cdots+m_d} \frac{s_l}{F}\frac{\partial F}{\partial s_l} \\
&= \left[s^{m_1-i_1} \cdots s^{m_l-i_l-1} \cdots s^{m_d-i_d}\right] \left(F(\underline{s})\right)^{m_1+\cdots+m_d} \frac{1}{F}\frac{\partial F}{\partial s_l} \\
&= \left[s^{m_1-i_1} \cdots s^{m_l-i_l-1} \cdots s^{m_d-i_d}\right] \left(F(\underline{s})\right)^{m_1+\cdots+m_d-1} \frac{\partial F}{\partial s_l} \\
&= \left[s^{m_1-i_1} \cdots s^{m_l-i_l-1} \cdots s^{m_d-i_d}\right] \frac{1}{m_1+\cdots+m_d} \frac{\partial \left(F(\underline{s})\right)^{m_1+\cdots+m_d}}{\partial s_l} \\
&= \left[s^{m_1-i_1} \cdots s^{m_l-i_l} \cdots s^{m_d-i_d}\right] \frac{m_l - i_l}{m_1+\cdots+m_d} \left(F(\underline{s})\right)^{m_1+\cdots+m_d} \\
&= \frac{m_l - i_l}{m_1+\cdots+m_d} P\left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d)\right)
\end{aligned}$$

Therefore

$$
\begin{aligned}
&P\left(\underline{S}=\underline{m} \mid \underline{Z}_0=\underline{i}\right) \\
&=\left(1-\sum_{l=1}^{d} \frac{m_l-i_l}{m_1+\cdots+m_d}\right) P\left(\underline{Z}_1=(m_1-i_1, \cdots, m_d-i_d) \mid \underline{Z}_0=(m_1, \cdots, m_d)\right) \\
&=\frac{i_1+\cdots+i_d}{m_1+\cdots+m_d} P\left(\underline{Z}_1=(m_1-i_1, \cdots, m_d-i_d) \mid \underline{Z}_0=(m_1, \cdots, m_d)\right)
\end{aligned}
$$

□

Theorem 4.5. *If the branching process starts off with i_1 individuals of type 1, i_2 of type 2, $\cdots$, i_d of type d, and suppose that the distributions of children are in the sense that*

$$f^{(\kappa)}(s_1, \cdots, s_d)=[F(s_1, \cdots, s_d)]^{p_\kappa} \quad \kappa=1,2, \cdots, d$$

with $f^{(\kappa)}(0, \cdots, 0) \neq 0$. Then the probability that the whole process will contain precisely m_1 individuals of type 1, $\cdots$, m_d of type d, is given by

$$P(\underline{S}=\underline{m} \mid \underline{Z}_0=\underline{i})=\frac{p_1 i_1+\cdots+p_d i_d}{p_1 m_1+\cdots+p_d m_d} P\left(\underline{Z}_1=(m_1-i_1, \cdots, m_d-i_d) \mid \underline{Z}_0=(m_1, \cdots, m_d)\right)$$

Proof of Theorem 4.5. Note from Theorem 3.4

$$
\begin{aligned}
\left\| \delta_\mu^\nu - \frac{s_\mu}{f^{(\mu)}(\underline{s})} \frac{\partial f^{(\mu)}(\underline{s})}{\partial s_\nu} \right\| &= \left\| \begin{matrix} 1 - \frac{s_1}{F^{p_1}} \frac{\partial F^{p_1}}{\partial s_1} & -\frac{s_1}{F^{p_1}} \frac{\partial F^{p_1}}{\partial s_2} & \cdots & -\frac{s_1}{F^{p_1}} \frac{\partial F^{p_1}}{\partial s_d} \\ -\frac{s_2}{F^{p_2}} \frac{\partial F^{p_2}}{\partial s_1} & 1 - \frac{s_2}{F^{p_2}} \frac{\partial F^{p_2}}{\partial s_2} & \cdots & -\frac{s_2}{F^{p_2}} \frac{\partial F^{p_2}}{\partial s_d} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{s_d}{F^{p_d}} \frac{\partial F^{p_d}}{\partial s_1} & -\frac{s_d}{F^{p_d}} \frac{\partial F^{p_d}}{\partial s_2} & \cdots & 1 - \frac{s_d}{F^{p_d}} \frac{\partial F^{p_d}}{\partial s_d} \end{matrix} \right\| \\
&= \left\| \begin{matrix} 1 - \frac{p_1 s_1}{F} \frac{\partial F}{\partial s_1} & -\frac{p_1 s_1}{F} \frac{\partial F}{\partial s_2} & \cdots & -\frac{p_1 s_1}{F} \frac{\partial F}{\partial s_d} \\ -\frac{p_2 s_2}{F} \frac{\partial F}{\partial s_1} & 1 - \frac{p_2 s_2}{F} \frac{\partial F}{\partial s_2} & \cdots & -\frac{p_2 s_2}{F} \frac{\partial F}{\partial s_d} \\ \cdots & \cdots & \cdots & \cdots \\ -\frac{p_p s_d}{F} \frac{\partial F}{\partial s_1} & -\frac{p_d s_d}{F} \frac{\partial F}{\partial s_2} & \cdots & 1 - \frac{p_d s_d}{F} \frac{\partial F}{\partial s_d} \end{matrix} \right\| \\
&= 1 - \sum_{l=1}^{d} \frac{p_l s_l}{F} \frac{\partial F}{\partial s_l}
\end{aligned}
$$

As in the proof of Theorem 4.1, we have

$$
\begin{aligned}
&[s^{m_1 - i_1} \cdots s^{m_d - i_d}] \, (F(\underline{s}))^{p_1 m_1 + \cdots + p_d m_d} \frac{p_l s_l}{F} \frac{\partial F}{\partial s_l} \\
&= [s^{m_1 - i_1} \cdots s^{m_l - i_l - 1} \cdots s^{m_d - i_d}] p_l \, (F(\underline{s}))^{p_1 m_1 + \cdots + p_d m_d - 1} \frac{\partial F}{\partial s_l} \\
&= [s^{m_1 - i_1} \cdots s^{m_l - i_l - 1} \cdots s^{m_d - i_d}] \frac{p_l}{p_1 m_1 + \cdots + p_d m_d} \frac{\partial \, (F(\underline{s}))^{p_1 m_1 + \cdots + p_d m_d}}{\partial s_l} \\
&= [s^{m_1 - i_1} \cdots s^{m_l - i_l} \cdots s^{m_d - i_d}] \frac{p_l \, (m_l - i_l)}{p_1 m_1 + \cdots + p_d m_d} \, (F(\underline{s}))^{p_1 m_1 + \cdots + p_d m_d} \\
&= \frac{p_l \, (m_l - i_l)}{p_1 m_1 + \cdots + p_d m_d} P \left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d) \right)
\end{aligned}
$$

Therefore

$$
\begin{aligned}
&P \left(\underline{S} = \underline{m} \mid \underline{Z}_0 = \underline{i} \right) \\
&= \left(1 - \sum_{l=1}^{d} \frac{p_l \, (m_l - i_l)}{p_1 m_1 + \cdots + p_d m_d} \right) P \left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d) \right) \\
&= \frac{p_1 i_1 + \cdots + p_d i_d}{p_1 m_1 + \cdots + p_d m_d} P \left(\underline{Z}_1 = (m_1 - i_1, \cdots, m_d - i_d) \mid \underline{Z}_0 = (m_1, \cdots, m_d) \right)
\end{aligned}
$$

□

Chapter 5

Independent Poisson Case

In this chapter, we consider different offspring probability generating function ($p.g.f$) but each type produces Poisson offspring.

Definition 5.1. Let a random variable X take on values 0, 1, $\cdots$ with probabilities

$$P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!}$$

where $\lambda > 0$. We say that X has a Poisson distribution with mean λ. Its generating function is given by

$$f_X(s) = e^{\lambda(s-1)}$$

Lemma 5.1. *Suppose that the distributions of children are independent Poisson with* $p.g.f$

$$\begin{aligned} f^{(1)}(s_1, s_2) &= e^{a_{11}(s_1-1)} e^{a_{12}(s_2-1)} \\ f^{(2)}(s_1, s_2) &= e^{a_{21}(s_1-1)} e^{a_{22}(s_2-1)} \end{aligned}$$

Then

$$\begin{aligned}\left(f^{(1)}(s_1,s_2)\right)^{m_1} &= e^{m_1a_{11}(s_1-1)}e^{m_1a_{12}(s_2-1)}\\ \left(f^{(2)}(s_1,s_2)\right)^{m_2} &= e^{m_2a_{21}(s_1-1)}e^{m_2a_{22}(s_2-1)}\end{aligned}$$

And so

$$\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2} = e^{(m_1a_{11}+m_2a_{21})(s_1-1)}e^{(m_1a_{12}+m_2a_{22})(s_2-1)} \tag{5.1}$$

where $e^{(m_1a_{11}+m_2a_{21})(s_1-1)}$ *and* $e^{(m_1a_{12}+m_2a_{22})(s_2-1)}$ *are probability generating functions of Poisson distribution with mean* $m_1a_{11}+m_2a_{21}$ *and* $m_1a_{12}+m_2a_{22}$*, respectively.*

Lemma 5.2. *For Poisson:* $\underline{Z}_0=(m_1,0)$*, the distribution of* $\left(Z_1^{(1)},Z_1^{(2)}\right)$ *is independent Poisson with means* (m_1a_{11},m_1a_{12})*. For Poisson:* $\underline{Z}_0=(0,m_2)$*, distribution of* $\left(Z_1^{(1)},Z_1^{(2)}\right)$ *is independent Poisson with means* (m_2a_{21},m_2a_{22})*. Therefore, for Poisson:* $\underline{Z}_0=(m_1,m_2)$*, distribution of* $\left(Z_1^{(1)},Z_1^{(2)}\right)$ *is independent Poisson with means* $(m_1a_{11}+m_2a_{21},m_1a_{12}+m_2a_{22})$*, and then*

$$\begin{aligned}&P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid \underline{Z}_0=(m_1,m_2)\right)\\ &=\frac{(m_1a_{11}+m_2a_{21})^{m_1-i_1}\,e^{-(m_1a_{11}+m_2a_{21})}}{(m_1-i_1)!}\times\frac{(m_1a_{12}+m_2a_{22})^{m_2-i_2}\,e^{-(m_1a_{12}+m_2a_{22})}}{(m_2-i_2)!}\end{aligned} \tag{5.2}$$

Proposition 5.3. *For Poisson offspring distribution and d=2, we have*

$$\begin{aligned}P(\underline{S}=\underline{m}\mid\underline{Z}_0=\underline{i}) &= \frac{(m_1a_{11}+m_2a_{21})\,a_{12}i_1+(m_1a_{12}+m_2a_{22})\,a_{21}i_2+(a_{11}a_{22}-a_{12}a_{21})i_1i_2}{(m_1a_{11}+m_2a_{21})\,(m_1a_{12}+m_2a_{22})}\\ &\quad\times P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)\end{aligned} \tag{5.3}$$

Proof. Since

$$\frac{s_1}{f^{(1)}}\frac{\partial f^{(1)}}{\partial s_1} = \frac{s_1}{f^{(1)}}a_{11}f^{(1)} = a_{11}s_1$$
$$\frac{s_1}{f^{(1)}}\frac{\partial f^{(1)}}{\partial s_2} = \frac{s_1}{f^{(1)}}a_{12}f^{(1)} = a_{12}s_1$$
$$\frac{s_2}{f^{(2)}}\frac{\partial f^{(2)}}{\partial s_1} = \frac{s_2}{f^{(2)}}a_{21}f^{(2)} = a_{21}s_2$$
$$\frac{s_2}{f^{(2)}}\frac{\partial f^{(2)}}{\partial s_2} = \frac{s_2}{f^{(2)}}a_{22}f^{(2)} = a_{22}s_2$$

And so

$$\left\|\delta_\mu^\nu - \frac{s_\mu}{f^{(\mu)}(\underline{s})}\frac{\partial f^{(\mu)}(\underline{s})}{\partial s_\nu}\right\| = \left\|\begin{matrix} 1-a_{11}s_1 & -a_{12}s_1 \\ -a_{21}s_2 & 1-a_{22}s_2 \end{matrix}\right\| \tag{5.4}$$
$$= 1 - a_{11}s_1 - a_{22}s_2 + (a_{11}a_{22} - a_{12}a_{21})s_1s_2$$

By Proposition 3.5, we have

$$\begin{aligned}
&P(\underline{S} = \underline{m} \mid \underline{Z}_0 = \underline{i}) \\
&= [s_1^{m_1-i_1}s_2^{m_2-i_2}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2} \\
&- a_{11}[s_1^{m_1-i_1-1}s_2^{m_2-i_2}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2} \\
&- a_{22}[s_1^{m_1-i_1}s_2^{m_2-i_2-1}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2} \\
&+ (a_{11}a_{22} - a_{12}a_{21})[s_1^{m_1-i_1-1}s_2^{m_2-i_2-1}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2}
\end{aligned} \tag{5.5}$$

since

$$[s_1^{m_1-i_1-1}s_2^{m_2-i_2}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2}$$
$$=\frac{m_1-i_1}{m_1a_{11}+m_2a_{21}}P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)$$
$$[s_1^{m_1-i_1}s_2^{m_2-i_2-1}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2}$$
$$=\frac{m_2-i_2}{m_1a_{12}+m_2a_{22}}P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)$$
$$[s_1^{m_1-i_1-1}s_2^{m_2-i_2-1}]\left(f^{(1)}(s_1,s_2)\right)^{m_1}\left(f^{(2)}(s_1,s_2)\right)^{m_2}$$
$$=\frac{(m_1-i_1)(m_2-i_2)}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)$$

Then

$$1-\frac{(m_1-i_1)a_{11}}{m_1a_{11}+m_2a_{21}}-\frac{(m_2-i_2)a_{22}}{m_1a_{12}+m_2a_{22}}+\frac{(m_1-i_1)(m_2-i_2)}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}(a_{11}a_{22}-a_{12}a_{21})$$
$$=1-\frac{m_1^2a_{11}a_{12}+m_1m_2a_{11}a_{12}-m_1a_{11}a_{12}i_1-m_2a_{11}a_{22}i_1}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}$$
$$-\frac{m_1m_2a_{11}a_{22}+m_2^2a_{21}a_{22}-m_1a_{11}a_{22}i_2-m_2a_{21}a_{22}i_2}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}$$
$$+\frac{m_1m_2a_{11}a_{22}-m_1m_2a_{12}a_{21}-m_1a_{11}a_{22}i_2}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}$$
$$+\frac{m_1a_{12}a_{21}i_2-m_2a_{11}a_{22}i_1+m_2a_{12}a_{21}i_1+(a_{11}a_{22}-a_{12}a_{21})i_1i_2}{(m_1a_{11}+m_2a_{21})(m_1a_{12}+m_2a_{22})}$$
$$=\frac{(m_1a_{11}+m_2a_{21})\,a_{12}i_1+(m_1a_{12}+m_2a_{22})\,a_{21}i_2+(a_{11}a_{22}-a_{12}a_{21})i_1i_2}{(m_1a_{11}+m_2a_{21})\,(m_1a_{12}+m_2a_{22})}$$

Therefore

$$P(\underline{S}=\underline{m}\mid\underline{Z}_0=\underline{i})=\frac{(m_1a_{11}+m_2a_{21})\,a_{12}i_1+(m_1a_{12}+m_2a_{22})\,a_{21}i_2+(a_{11}a_{22}-a_{12}a_{21})i_1i_2}{(m_1a_{11}+m_2a_{21})\,(m_1a_{12}+m_2a_{22})}$$
$$\times P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)$$

□

If $a_{11}=a_{21}$, $a_{12}=a_{22}$, then reduce to

$$P(\underline{S}=\underline{m}\mid\underline{Z}_0=\underline{i})=\frac{i_1+i_2}{m_1+m_2}P\left(\underline{Z}_1=(m_1-i_1,m_2-i_2)\mid\underline{Z}_0=(m_1,m_2)\right)$$

Chapter 6

Special Case

We let

$$\{Z_n^{(i)} : n = 0, 1, \cdots\}$$

denote the branching process with initial i particles. In general, write $Z_n^{(1)} = Z_n$.

Suppose that we have a branching process $\{Z_n\}$ with birth law

$$f(s) = E\left(s^{Z_1}|Z_0 = 1\right) = 1 - p - ps$$

which is Bernoulli with

$$\begin{aligned} P(Z_1 = 1) &= p \\ P(Z_1 = 0) &= 1 - p \end{aligned}$$

In this sense,

$$P(Z_1 = m - k \mid Z_0 = m) = \binom{m}{m-k} p^{m-k}(1-p)^k$$

Note that f^m is the generating function of a binomial distribution $B(m, p)$.

1. Given $k = 1$

   The total progeny S has a geometric distribution $G(p)$ with support on the set $\{1, 2, 3, \cdots\}$, and probability mass function (pmf) given by

   $$P(S = m|Z_0 = 1) = (1-p)p^{m-1} \quad m = 1, 2, \cdots$$

   and generating function

   $$f_S(s) = \sum_{m=1}^{\infty}(1-p)p^{m-1}s^m = \frac{1}{p}\frac{1-p}{1-ps} \propto \frac{1-p}{1-ps}$$

   Thus

   $$P(S = m) = \frac{p}{m}P(Y = m-1) \qquad (6.1)$$

   where $S \sim G(p),\ Y \sim B(m,p)$

2. Given $k \geq 2$

   $$S = X_1 + \cdots + X_k$$

   where $X_i, i = 1, \cdots, k$, has a geometric distribution with support on the set $\{1, 2, 3, \cdots\}$. Therefore, the total progeny S has a negative binomial distribution

   $$f_S(s) = \frac{1}{p^k}\left(\frac{1-p}{1-ps}\right)^k$$

   Thus

   $$P(S = m) = \frac{kp^k}{m}P(Y = m-k) \qquad (6.2)$$

   where $S \sim NB(k,p),\ Y \sim B(m,p)$

Appendix A

Notation Index

Symbol	Description
$\underline{a}$	vector $\underline{a} = (a_1, \cdots, a_d)$
$a_{\underline{i}}$	$a_{i_1 \cdots i_d}$
$\underline{z}^{\underline{n}}$	$z_1^{n_1} \cdots z_d^{n_d}$
$\underline{e}_i$	the vector whose i^{th} component is 1 and whose other components are 0
$\|\cdot\|$	determinant of matrix
$p.g.f$	probability generating function
$E(X)$	expectation of random variable X
Z_n	single type branching process
$\underline{Z}_n$	multitype type branching process
$\mathbb{R}$	real field
$\mathbb{C}$	complex field

Bibliography

[1] A. M. Odlyzko, *Asymptotic Enumeration Methods*, in Handbook of Combinatorics, vol. 2, R. L. Graham, M. Groetschel, and L. Lovasz, eds., Elsevier, 1995, pp. 1063–1229

[2] C. Genovese, *Lecture Notes:Notes on Generating Functions*, Retrieved Sep.3,2007 from
http : //lib.stat.cmu.edu/ ∼ genovese/class/iprob − S06/notes/generating − functions.pdf

[3] D. Merlini, R.Sprugnoli and M.C.Verri, *Lagrange Inversion:When and How*, Acta. Appl. Math., **94**(2006):233-249.

[4] D. Stebila, *Combinatorial Enumeration*, In Gregory V. Bard, ed., Vanquishing the Tyranny of Calculus (working title), in progress, 2006. Retrieved Sep.3,2007 from
http : //www.douglas.stebila.ca/files/research/papers/Ste06vtc.pdf

[5] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Fourth Edition, reprinted), Cambridge, England, University Press, 1946.

[6] F. M. Hoppe, *Branching Processes and The Effect of Parlaying Bets on Lottery Odds*, Optimal Play: Mathematical Studies of Games and Gambling, Institute for the Study of Gambling and Commercial Gaming, University of Nevada, in press (2007).

[7] I. J. Good, *The Joint Distribution of The Sizes of The Generations in a Cascade Process*, Proc. Cambridge Philos. Soc.,**51**(1955):240–242.

[8] I. J. Good, *Generalizations to Several Variables of Lagrange's Expansion, with Applications to Stochastic Processes*, Proc. Cambridge Philos. Soc.,**56**(1960):367–380.

[9] I. J. Good, *The Lagrange Distributions and Branching Processes*, SIAM Journal on Applied Mathematics,**28**(1975):270–275.

[10] I. M. Gessel, *A Combinatorial Proof of the Multivariable Lagrange Inversion Formula*, J. Combin. Theory Ser. A45 (1987), 178–195.

[11] J. Chaumat and A.-M. Chollet, *On Composite Formal Power Series*, Trans. Amer. Math. Soc. **353**(2001), no. 4, 1691–1703.

[12] J. A. Gallian, *Contemporary Abstract Algebra*, 5th ed., Houghton Mifflin Company, 2002.

[13] J. G. Wendel, *Left Continuous Random Walk and The Lagrange Expansion*, Amer. Math. Monthly, **82**(1975):494–498.

[14] J. S. Vitter and Ph. Flajolet, *"Average-Case Analysis of Algorithms and Data Structures,"* , Chapter 9 in Handbook of Theoretical Computer Science, Volume A: Algorithms and Complexity (edited by J. van Leeuwen), Elsevier and M.I.T. Press, 1990, 431–524.

[15] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer-Verlag, Berlin, 1972.

[16] M. Dwass, *The Total Progeny in a Branching Process and a Related Random Walk*, J. Appl. Prob., **6**(1969):682–686.

[17] R. Otter, *The Multiplicative process*, Ann. Math. Statist. **20**(1949):206–224

[18] R. Sprugnoli, *An Introduction to Mathematical Methods in Combinatorics*, Retrieved Sep.3,2007 from *http : //www.dsi.unifi.it/ resp/Handbook.pdf*

[19] I. R. Shafarevich, *Basic Notions of Algebra*,2nd Edition, Springer-Verlag, 1997

[20] S. Roman, *Advanced Linear Algebra*, Graduate Texts in Mathematics, vol. 135, Springer-Verlag, New York, 1992.

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

[22] X. Gan and N. Knox, *On Composite Formal Power Series* ,International Journal of Mathematics and Mathematical Sciences, **30** (2002): Issue 12, 761–770