The Distribution of the Length of the Longest Increasing Subsequence in Random Permutations of Arbitrary Multi-sets

by

Ayat Al-Meanazel

A thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics
University of Manitoba
Winnipeg

Abstract

The distribution theory of runs and patterns has a long and rich history. In this dissertation we study the distribution of some run-related statistics in sequences and random permutations of arbitrary multi-sets. Using the finite Markov chain imbedding technique (FMCI), which was proposed by Fu and Koutras (1994), we proposed an alternative method to calculate the exact distribution of the total number of adjacent increasing and adjacent consecutive increasing subsequences in sequences.

Fu and Hsieh (2015) obtained the exact distribution of the length of the longest increasing subsequence in random permutations. To the best of our knowledge, little or no work has been done on the exact distribution of the length of the longest increasing subsequence in random permutations of arbitrary multi-sets. Here we obtained the exact distribution of the length of the longest increasing subsequence in random permutations of arbitrary multi-sets. We also obtain the the exact distribution of the length of the longest increasing subsequence for the set of all permutations of length N generated from $\{1, 2, \ldots, n\}$.

Acknowledgment

Glory and praise be to Allah Almighty for His immense blessings and help provided to me throughout my life including granting me the health, and strength to fulfil my dream to complete my PhD degree.

I would like to thank my supervisor Dr. Brad Johnson for his support, patient guidance, encouragement, and advice he has provided throughout my study. I also thank my committee members, Dr. James Fu, and Dr. David Gunderson for their recommendations to improve the thesis, and Dr. Markos Koutras for accepting to be my External Examiner. I would also like to thank to Dr. A. Thavaneswaran, and Dr. Saumen Mandal for their support and encouragement during my time as their student.

I must express my gratitude to my companions on this long and exhausted journey, who experienced all of the ups and downs of my study. Without their unconditional support and love I would not be here today. My best friend and beloved husband Sadeq, who sacrificed as much as I did. My children Hassan, Eba'a, and Meral, who lighten my life. And my parents Nadwa and Taisir, you are the reason behind every success in my life. I would also thank my family and family in law for their prayers and support.

Finally, I would like to thank Al Al-Bayt University for providing me with the finical support to complete my PhD.

*This thesis is dedicated to
my beloved family*

Contents

Chapter 1

Introduction

The distribution theory of runs has a rich history. As early as 1738, De Moivre (1738) studied the probability of getting a success run of length r or more in n trials in his book *The Doctrine of Chances*. Many researchers have discussed this problem further; for more details, see Mood (1940).

In the 1940's many researchers were interested in the concept of runs and patterns in Bernoulli trials, but the main focus was on the conditional distributions of runs given the total number of successes (see, for example, Stevens, 1939). Even though little research was done on the exact distribution, there were some useful approximation formulas for the distributions of runs developed in the period between 1950 and 1970 (see, for example, Walsh, 1962; Gibbons, 1985).

When interest in run-related statistics began in the nineteenth century, the tools used to study the distributions of run-related statistics were combinatorial methods. Due to the complexity in the combinatorial

calculations, the exact distribution of many run-related statistics remains unknown. For example, in the cases where the study involves elements of more than two types and/or they are non-iid trials, the formula tend to be more complicated and difficult to compute. Much of the research in the nineteenth century focused on finding the conditional distribution of the number of success runs given the number of successes in the iid Bernoulli trials (see, for example, Wald and Wolfowitz, 1940). Despite the complexity of this approach, it gave some interesting results. For example, Hirano (1986), Philippou and Makri (1986), and Godbole (1990) all were able to find the exact distribution for non-overlapping consecutive success runs in iid Bernoulli trials.

Many researchers started to realize the complexity of finding the exact distribution of run-related statistics using combinatorial methods (see, for example, Aki and Hirano, 1994). It did not take long until they revealed the recurrence relations hidden within the sequence of trials, and by identifying these relations they were able to use probability generating functions to find the exact distribution of the run-related statistic. After deriving the probability generating function, it could be differentiated the desired number of times to yield the probability distribution function. Ling (1988) gave recurrence relations for the probability generating function of the number of overlapping consecutive k successes until the n-th trial for iid trials and an extension for non-iid trials was proposed by Chryssaphinou et al. (1993).

Both the combinatorial approach and the probability generating func-

2

tion approach shared the same difficulties of complicated calculation and the complexity of the results. These methods become increasingly difficult when studying more complicated patterns. That left the exact distribution of many run-related statistics unknown. Instead, researchers used approximation methods, such as the Chen-Stein method and compound Poisson approximations, to help them with these cases (see, for example, Chen, 1975). These approximation methods have their own disadvantages, such as the difficulty in finding the bounds of the error and defining the distribution parameters. Boutsikas and Koutras (2000) presented an upper bound for the approximation error and offered a simple formulae for the distribution parameters.

Fu and Koutras (1994) introduced an entirely different approach based on the Chapman-Kolmogorov equation and the ability to imbed a random variable into a finite Markov chain. They proposed the "finite Markov chain imbedding technique" (FMCI) to solve some run-related statistic problems via transition matrices of the induced Markov chain. This approach has been used in many areas, such as reliability and quality control (see, for example, Fu and Hu, 1987; Chao and Fu, 1989, 1991), and random permutation problems (see, for example, Fu, 1995; Fu et al., 1999; Johnson and Fu, 2000).

Brook and Evans (1972) used the Markov chain representation for calculating the distribution of the run-length of a Cumulative Sum control chart (CUSUM). Chao and Lin (1984) used the Markov chain imbedding technique to study the reliability of "consecutive-k-out-of-n:F" systems (a

3

system of n ordered components such that the system fails if, and only if, k-consecutive components fail). They imbedded the general "consecutive-k-out-of-n:F " system into a Markov chain with 2^k states. Fu (1986) was able to reduce the size of the state space of the general "consecutive-k-out-of-n:F " system to $(k-1)$. Later, Fu and Hu (1987), and Chao and Fu (1989, 1991) proposed a formula for the general case of s-independent elements. Fu and Koutras (1994) formalized the finite Markov chain imbedding technique by adopting the ideas of Chao and Lin (1984), Fu (1986), Fu and Hu (1987), and Chao and Fu (1989, 1991) papers on system reliability.

The finite Markov chain imbedding technique has been successfully applied in different cases and models and it gave simple and clear expressions that were easy to work with. For example, Fu and Koutras (1994) found the exact distributions of many run-related statistics with the finite Markov chain imbedding technique.

Fix N, n and let $[\boldsymbol{d}_n]$ be a partition of N into n parts so that $[\boldsymbol{d}_n] = [d_1, d_2, \ldots, d_n]$, where the d_i's are non-negative integers, and let $N = d_1 + d_2 + \cdots + d_n$. The naturally ordered multi-set with specification $[\boldsymbol{d}_n]$ is the set $D([\boldsymbol{d}_n]) = \{1, \ldots, 1, 2, \ldots, 2, \ldots, n, \ldots, n\}$, where there are d_i copies of the integer "i" for each $i = 1, 2, \ldots, n$. Let $\prod[\boldsymbol{d}_n]$ denote the set of all permutations of $D([\boldsymbol{d}_n])$. For example, for $N = 7$, $n = 4$ and $[d_4] = [2, 1, 1, 3]$, we have $D([d_4]) = \{1, 1, 2, 3, 4, 4, 4\}$. Note that, with $[\boldsymbol{1}_n] = [1, 1, \ldots, 1]$, the set $\prod([\boldsymbol{1}_n])$ is the usual symmetric group on $\{1, 2, \ldots, n\}$.

Every permutation $\pi \in \prod[\boldsymbol{d}_n]$ can be obtained by inserting the N integers from $D([\boldsymbol{d}_n])$ one by one into gaps between previously inserted integers (including the two end gaps), starting by inserting the d_1 integers "1", followed by inserting the d_2 integers "2", and continuing this insertion procedure until the last d_n integers "n" have been inserted. In general, for $t = 1, 2, \ldots, N$, the t-th element inserted is the $c(t)$-th copy of the integer "$k(t)$", where

$$k(t) = \min\{k \in \mathbb{N} \mid \sum_{r=1}^{k} d_r \geq t\}, \text{ and } c(t) = t - \sum_{r=1}^{k(t)-1} d_r.$$

If the insertion positions are randomly chosen, then the sequence of insertions can be viewed as a sequence of trials. The above insertion procedure provides a way of studying the distributions of runs and patterns on random permutations via the finite Markov chain imbedding technique.

The number of random permutations of n elements with certain patterns has a rich history in mathematics, statistics and computer science. For example, many researchers have studied the number of random permutations with exactly k increasing subsequences of length ℓ (see, for example, Carlitz et al., 1964; Roselle, 1968; Dillon and Roselle, 1969; Dwass, 1973; Tanny et al., 1973; Tanny, 1976; Reilly and Tanny, 1979; Johnson, 2001). For a given permutation $\pi \in \prod[\boldsymbol{d}_n]$ there are several types of increasing subsequences, such as:

- Increasing subsequence (IS): A permutation $\pi \in \prod[\boldsymbol{d}_n]$ contains an increasing subsequence of length ℓ if there exists $\{e_1, \ldots, e_\ell\} \subseteq$

$\{1, \ldots, N\}$, such that, $e_1 < e_2 < \cdots < e_\ell$ and $\pi(e_1) < \pi(e_2) < \cdots < \pi(e_\ell)$.

- Adjacent increasing subsequence (AIS): A permutation $\pi \in \prod[\boldsymbol{d}_n]$ contains an adjacent increasing subsequence of length ℓ if, for some $i = 1, 2, \ldots, N - \ell + 1$, $\pi(i) < \pi(i+1) < \cdots < \pi(i + \ell - 1)$.

- Adjacent consecutive increasing subsequence (ACIS): A permutation $\pi \in \prod[\boldsymbol{d}_n]$ contains an adjacent consecutive increasing subsequence of length ℓ if, for some $i = 1, 2, \ldots, N - \ell + 1$, and for all $k = i, \ldots, i + \ell - 1$, $\pi(k) = \pi(k - 1) + 1$.

An adjacent increasing subsequence of length ℓ is sometimes referred to as a rise of length ℓ (see, for example, Roselle, 1968), and an adjacent consecutive increasing subsequence of length ℓ is sometimes referred to as a succession of length ℓ (see, for example, Fu, 1995). It is clear that an adjacent consecutive increasing subsequence of length ℓ is also an adjacent increasing subsequence of length ℓ, and an adjacent increasing subsequence of length ℓ is also an increasing subsequence of length ℓ. Thus for any permutation $\pi \in \prod[\boldsymbol{d}_n]$ the total number of adjacent consecutive increasing subsequences of length ℓ is less than or equal to the total number of adjacent increasing subsequences of length ℓ, which is less than or equal to the total number of increasing subsequences of length ℓ. There are a number of problems of interest with these run-related statistics in π, for example:

- What is the probability that a random permutation $\pi \in \prod[\boldsymbol{d}_n]$ has

6

exactly k adjacent consecutive increasing subsequences of length ℓ?

- What is the probability that a random permutation $\pi \in \prod[\boldsymbol{d}_n]$ has exactly k adjacent increasing subsequences of length ℓ?

- What is the probability that a random permutation $\pi \in \prod[\boldsymbol{d}_n]$ has exactly k increasing subsequences of length ℓ?

- What is the probability that the longest increasing subsequence in a random permutation $\pi \in \prod[\boldsymbol{d}_n]$ is of length ℓ?

We studied the first two cases for sequences, as well as the probability that the longest increasing subsequence in a random permutation $\pi \in \prod[\boldsymbol{d}_n]$ is of length ℓ.

Let $A([\boldsymbol{d}_n], \ell, k)$ denote the total number of permutations in $\prod[\boldsymbol{d}_n]$ with exactly k adjacent increasing subsequences of length ℓ. The most studied numbers associated with $A([\boldsymbol{d}_n], \ell, k)$ are the Eulerian numbers, which are given by $A([\boldsymbol{1}_n], 2, k)$ and the Simon Newcomb numbers given by $A([\boldsymbol{d}_n], 2, k)$. The Eulerian and Simon Newcomb numbers have been studied extensively (see, for example, Carlitz et al., 1964; Dillon and Roselle, 1969; Tanny et al., 1973; Carlitz, 1974; Nicolas, 1991; Giladi and Keller, 1994; Harris and Park, 1994) and references therein. Fu et al. (1999) used the finite Markov chain imbedding technique to find the exact distribution of the Eulerian and Simon Newcomb numbers. To the best of our knowledge, little or no work has been done on $A([\boldsymbol{d}_n], \ell, k)$ for $\ell > 2$.

Let $B([\boldsymbol{d}_n], \ell, k)$ denote the total number of permutations in $\prod[\boldsymbol{d}_n]$ with exactly k adjacent consecutive increasing subsequences of length ℓ.

Whitworth (1901) was the first to discuss the number of permutations in $\prod[\mathbf{1}_n]$ with no increasing subsequences of length 2. The number of permutations in $\prod[\mathbf{1}_n]$ with no increasing and/or decreasing subsequences of length 2 is examined by Wolfowitz (1942, 1944), Kaplansky (1944, 1945), and Abramson and Moser (1967). Many researchers have studied $B([\mathbf{1}_n], \ell, k)$ (see, for example, Roselle, 1968; Dwass, 1973; Roselle et al., 1974; Tanny, 1976; Dymacek and Roselle, 1978; Reilly and Tanny, 1979). Permutations in $\prod[\mathbf{1}_n]$ with increasing subsequences of length 3 are studied by Riordan (1945) and the more general case of increasing subsequences of length ℓ has been examined by Abramson and Moser (1967). Jackson and Goulden (1980, 1981) developed a general theory on sequence enumeration, which allows us to obtain the generating functions for a wide class of combinatorial problems. For example, when this theory is applied to study $B([\mathbf{1}_n], \ell, k)$ the results lead to those contained in Jackson and Reilly (1976), and Jackson and Aleliunas (1977). Most recently, Johnson (2001) examined the distribution of $B([\mathbf{1}_n], \ell, k)$ based on the compositions of n elements, and proposed a new solution which required, at most, $k(n-k-\ell)$ summands.

The finite Markov chain imbedding technique has successfully been applied to $B([\mathbf{1}_n], \ell, k)$. A simple method based on the finite Markov chain imbedding technique is used by Fu (1995) to obtain the exact distribution of $B([\mathbf{1}_n], 2, k)$. His approach also gives a direct proof that the limiting distribution of $B([\mathbf{1}_n], \ell, k)$ is a Poisson distribution with parameter $\lambda = 1$. Johnson and Fu (2000) studied the general case of $B([\mathbf{1}_n], \ell, k)$ using the

8

finite Markov chain imbedding technique. Johnson (2002) obtained the exact distribution of $B([\boldsymbol{d}_n], 2, k)$ using the finite Markov chain imbedding technique. However, little work has been done on $B([\boldsymbol{d}_n], \ell, k)$ for $\ell > 2$.

Let $L_N([\boldsymbol{d}_n], \pi)$ denote the length of the longest increasing subsequence in a random permutation $\pi \in \Pi[\boldsymbol{d}_n]$. $L_n([\boldsymbol{1}_n], \pi)$ has been studied extensively (see, for example, Odlyzko and Rains, 2000; Deutsch et al., 2003; Romik, 2014). The expected value of the longest increasing subsequence $L_n([\boldsymbol{1}_n], \pi)$ in random permutation $\pi \in \Pi[\boldsymbol{1}_n]$ has also been studied extensively. There are interesting results using both Monte Carlo and exact computations that explore the finer structure of the distribution of $L_n([\boldsymbol{1}_n], \pi)$ (see, for example, Abramowitz and Stegun, 1964; Aldous and Diaconis, 1999; Baik et al., 1999).

Erdős and Szekeres (1935) showed that every permutation $\pi \in \Pi[\boldsymbol{1}_n]$ has either increasing or decreasing subsequence of length $\geq \sqrt{n-1} + 1$. Since then many researchers have been motivated to find the exact distribution of $L_n([\boldsymbol{1}_n], \pi)$. Schensted (1961) derived the Schensted correspondence, which has been a major tool in studying the distribution of $L_n([\boldsymbol{1}_n], \pi)$ in random permutations along with Young tableaus (see, for example, Baer and Brock, 1968; Odlyzko and Rains, 2000). Baer and Brock (1968) tabulated the exact distribution of $L_n([\boldsymbol{1}_n], \pi)$ for $n \leq 36$ using the Schensted correspondence along with Young tableaus. This result was extended by Odlyzko and Rains (2000) for $n \leq 120$, and they carried out Monte Carlo simulations for n up to 10^{10}.

Ulam (1961) used Monte Carlo simulations to suggest that

$$\lim_{n \to \infty} E(L_n([\mathbf{1}_n], \pi))/\sqrt{n} \equiv c \qquad (1.1)$$

and that $L_n([\mathbf{1}_n], \pi)$ is usually of the order $\sqrt{n}$. The problem of computing the value c was known as Ulam's problem. Several conjectures had been proposed about the value c (see, for example, Baer and Brock, 1968; Logan and Shepp, 1977; Pilpel, 1990; Aldous and Diaconis, 1995; Johansson, 1998). Vershik and Kerov (1977) showed that $c = 2$. Baik et al. (1999) obtained the limiting distribution of $L_n([\mathbf{1}_n], \pi)$ to be the Tracy-Widom distribution of the largest eigenvalue of a random Gaussian Unitary Ensemble (GUE) matrix.

Recently, Fu and Hsieh (2015) obtained the exact distribution of $L_n([\mathbf{1}_n], \pi)$ in random permutations using the finite Markov chain imbedding technique. To the best of our knowledge, no work has been done on the length of the longest increasing subsequence in random permutations of arbitrary multi-sets $L_N([\boldsymbol{d}_n], \pi)$.

This dissertation is organized as follows. Chapter 2 introduces some preliminary results regarding the finite Markov chain imbedding technique.

In Chapter 3 we find the exact distributions of the number of adjacent increasing subsequences, adjacent consecutive increasing subsequences. These results are general and can be applied to various selection probabilities. Numerical examples for some of these distributions are presented.

In Chapter 4 we first introduce the exact distribution of the length of the longest increasing subsequence in a random permutation of arbitrary multi-

sets (the conditional case). Then we obtained the exact distribution of the length of the longest increasing subsequence in the set of all permutations of length N generated from $\{1, 2, \ldots, n\}$ (the unconditional case). Finally, the summary of this thesis and future work are presented in Chapter 5.

Chapter 2

Finite Markov chain technique

2.1 Finite Markov chain

For $m < \infty$, let $\Omega = \{a_1, a_2, \ldots, a_m\}$ be a finite state space, and let $\{Y_t\}_{t \geq 0}$ be an ordered sequence of random variables defined on Ω. For ease of notation, we will frequently just write $\{Y_t\}$ for $\{Y_t\}_{t \geq 0}$.

Definition 2.1. *A sequence $\{Y_t\}$ of random variables is called a finite Markov chain if, for $i_t \in \Omega$,*

$$\Pr(Y_t = i_t \mid Y_{t-1} = i_{t-1}, \ldots, Y_0 = i_0) = \Pr(Y_t = i_t \mid Y_{t-1} = i_{t-1}). \quad (2.1)$$

In words, the sequence $\{Y_t\}$ is a finite Markov chain if the probability that the system enters the state i_t at time t depends only on the immediately preceding state i_{t-1} at time $t - 1$.

The one-step transition probabilities for the system at time t are the conditional probabilities given by

$$\Pr(Y_t = j \mid Y_{t-1} = i) \equiv p_{ij}(t), \quad i, j \in \Omega. \quad (2.2)$$

The transition probabilities $p_{ij}(t), 1 \le i, j \le m$ may be represented as an $m \times m$ matrix

$$M_t = (p_{ij}(t)) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1m}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}(t) & p_{m2}(t) & \cdots & p_{mm}(t) \end{bmatrix}.$$

For $t \ge 1$, the matrices M_t are called one-step transition probability matrices.

Definition 2.2. *A Markov chain $\{Y_t\}$ is homogenous if the transition probabilities are constant in time, i.e. for any $i, j \in \Omega$ and all $t = 1, 2, \ldots, n$, $\Pr(Y_t = j \mid Y_{t-1} = i) = p_{ij}$.*

2.2 Chapman-Kolmogorov equation

The n-step transition probabilities $\Pr(Y_{t+n} = j \mid Y_t = i) = p_{ij}^{(n)}(t)$ for a non-homogeneous Markov chain $\{Y_t\}$ can be obtained from the one-step transition probabilities using the Chapman-Kolmogorov equation. For example, let $n = 2$; then for $t \ge 2$

$$
\begin{aligned}
p_{ij}^{(2)}(t) &= \sum_k \Pr(Y_{t-1} = k \mid Y_{t-2} = i) \Pr(Y_t = j \mid Y_{t-1} = k) \\
&= \sum_k p_{ik}(t-1) p_{kj}(t)
\end{aligned}
\tag{2.3}
$$

which corresponds to summing over all possible intermediate states k in the transition from state i to state j.

14

If $\{Y_t\}$ is a homogeneous Markov chain, then equation (2.3) yields the two-step transition probabilities

$$
\begin{aligned}
p_{ij}^{(2)} &= \sum_k \Pr(Y_{t-1} = k \mid Y_{t-2} = i) \Pr(Y_t = j \mid Y_{t-1} = k) \\
&= \sum_k p_{ik} p_{kj},
\end{aligned}
\tag{2.4}
$$

which are independent of t.

For any subset $C \subseteq \Omega$ at time index n, any initial distribution $\zeta_0 = (\Pr(Y_0 = 1), \Pr(Y_0 = 2), \dots, \Pr(Y_0 = m))$, and a homogenous Markov chain $\{Y_t\}$, let

$$
\Pr(Y_n \in C \mid \zeta_0) = \zeta_0 M^n U^{\mathsf{T}}(C),
\tag{2.5}
$$

where $U^{\mathsf{T}}(C)$ denotes the transpose of $U(C)$, where $U(C) = \sum_{i \in C} e_i$, where e_i is a $1 \times m$ row vector with a 1 in the position corresponding to the i'th state and zeros elsewhere. Generally, if $\{Y_t\}$ is a non-homogeneous Markov chain, it can be shown (see, for example, Feller, 1968) that

$$
\Pr(Y_n \in C \mid \zeta_0) = \zeta_0 (\prod_{t=1}^n M_t) U^{\mathsf{T}}(C).
\tag{2.6}
$$

2.3 Finite Markov chain imbedding technique

A run denotes a sequence of consecutive successes or failures within a Bernoulli trials. There are several run statistics in a sequence of n Bernoulli

trials, such as the number of non-overlapping consecutive k successes $(N_{n,k})$, and the number of success runs of size greater than or equal to k $(G_{n,k})$. There are three type of general patterns of symbols: simple patterns; compound patterns; and series patterns. A simple pattern is composed of a specified sequence of k symbols. The length of the pattern is fixed, and the symbols in the pattern may be repeated. A compound pattern is a union of simple distinct patterns. Finally, a series pattern is an ordered sequence of distinct simple patterns.

Let $X_n(\Lambda)$ denote the number of occurrences of the pattern Λ in a sequence of n multi-state trials. The exact distribution of $X_n(\Lambda)$ depends on three factors: the structure of the pattern, the structure of the sequence of n multi-state trials, and the counting procedure (overlapping or non-overlapping counting).

Determining the exact distribution of $X_n(\Lambda)$ using the traditional methods, such as combinatorics, can be challenging for complex patterns and non-iid trials. Fu and Koutras (1994) introduced the finite Markov chain imbedding technique as an alternative to the traditional combinatorial methods. Let $\Gamma_n = \{0, 1, \ldots, n\}$ be an index set, and let $\Omega = \{a_1, a_2, \ldots, a_m\}$ be a finite state space.

Definition 2.3. *A non-negative integer-valued random variable $X_n(\Lambda)$ with support $S(X) = \{x : x = 0, 1, \ldots, l_n, \text{ and } \Pr(X = x) > 0\}$ is finite Markov chain imbeddable if:*

- *there exists a finite Markov chain $\{Y_t : t \in \Gamma_n\}$ defined on a finite*

state space Ω with initial probability vector ζ_0;

- *there exists a finite partition $\{C_x : x \in S(X)\}$ on the state space Ω, so that for every $x \in S(X)$,*

$$\Pr(X_n(\Lambda) = x) = \Pr(Y_n \in C_x \mid \zeta_0).$$

For $t = 1, 2, \ldots, n$, let M_t be the transition probability matrices of the induced finite Markov chain $\{Y_t\}$ defined on the state space Ω with initial probability $\zeta_0 = (\Pr(Y_0 = a_1), \Pr(Y_0 = a_2), \ldots, \Pr(Y_0 = a_m))$.

Theorem 2.1. *Let $X_n(\Lambda)$ be finite Markov chain imbeddable, and let $M_1, M_2, \ldots,$ denote the induced transition probability matrices. Then*

$$\Pr(X_n(\Lambda) = x) = \zeta_0(\prod_{t=1}^{n} M_t)U^{\mathsf{T}}(C_x), \tag{2.7}$$

The $k-$th moment $E(X_n^k(\Lambda))$, $k = 1, 2, \ldots,$ can be written as

$$E(X_n^k(\Lambda)) = \zeta_0(\prod_{t=1}^{n} M_t)V_k^{\mathsf{T}}, \tag{2.8}$$

where

$$V_k = \sum_{x=0}^{l_n} x^k U(C_x).$$

The probability generating function for the random variable $X_n(\Lambda)$ is given by

$$\varphi(s) = \zeta_0(\prod_{t=1}^{n} M_t)W^{\mathsf{T}}(s), \tag{2.9}$$

where

$$W(s) = \sum_{x=0}^{l_n} s^x U(C_x).$$

Koutras and Alexandrou (1995) introduced the Markov chain imbeddable variable of binomial type. They proposed a new model, which is based on Markov chain imbedding, to study of the exact distribution of variables related to success runs in a sequence of Bernoulli trials, scan statistics and success run statistics related to several well known urn models.

Definition 2.4. $X_n(\Lambda)$ *is called a Markov chain imbeddable variable of binomial type if:*

1. $X_n(\Lambda)$ *is Markov chain imbeddable,*

2. *for all* $x \in S(X)$*, the sets* $C_x = \{c_{x0}, c_{x1}, \ldots, c_{xs-1}\}$ *all have the same size* s*,*

3. *and* $\Pr(Y_t = c_{yj} \mid Y_{t-1} = c_{xi}) = 0$ *for all* $y \neq x, x + 1$*, and* $t = 1, \ldots, n\}.$

A Markov chain imbeddable variable of binomial type must visit its next state before jump directly to a higher state. Koutras and Alexandrou introduced two $s \times s$ transition probability matrices

$$A_t(x) = (\Pr(Y_t = c_{xj} \mid Y_{t-1} = c_{xi})),$$

and

$$B_t(x) = (\Pr(Y_t = c_{x+1j} \mid Y_{t-1} = c_{xi})).$$

Where the entries of $A_t(x)$ control the transitions within C_x, and the entries of $B_t(x)$ control the transitions between C_x and C_{x+1}. Note that, the sum $A_t(x) + B_t(x)$ is a stochastic matrix.

Using these transition probability matrices, they were able to develop the probability mass function of the Markov chain imbeddable variable of binomial type as follows:

$$\Pr(X_n(\Lambda) = x) = f_n(x)\mathbf{1}^\mathsf{T}, \tag{2.10}$$

where, for $0 \leq x \leq l_n$, $\mathbf{1}^\mathsf{T} = (1, 1, \ldots, 1)$ is a row vector and, for $t = 1, 2, \ldots, n$, the probability row vectors $f_t(x)$ are given by

$$f_t(x) = (\Pr(Y_t = c_{x0}), \Pr(Y_t = c_{x1}), \ldots, \Pr(Y_t = c_{xs-1}))$$

$$= \begin{cases} f_{t-1}(x)A_t(x) & \text{for } x = 0, \\ f_{t-1}(x)A_t(x) + f_{t-1}(x-1)B_t(x-1) & \text{for } 1 \leq x \leq l_n. \end{cases}$$

To illustrate the above definitions and to compare the Koutras and Alexandrou (1995) approach with the Fu and Koutras (1994) approach, consider a sequence of Bernoulli trials $Z_1, Z_2, \ldots$ with success probabilities $p_t = Pr(Z_t = 1)$, and failure probabilities $q_t = Pr(Z_t = 0) = 1 - p_t, t \geq 1$. To study the exact distribution of the number of non-overlapping consecutive k successes $N_{n,k}$, consider the state space

$$\Omega = \{(x, i) : x = 0, 1, \ldots, [n/k], \text{ and } i = 0, 1, \ldots, k-1\}.$$

Fu and Koutras (1994) defined a finite homogenous Markov chain $\{Y_t\}_{t=1}^n$ on Ω as follows:

$$Y_t = (N_{t,k}, E_t), \text{ for } 1 \leq t \leq n,$$

where E_t equals m modulo k, and m is the number of trailing successes counting backwards that exist in the sequence after the first t trials (*e.g.*,

$SFFS \ldots F \overbrace{SS \ldots S}^{m}$). The transition probability matrix M_t of the induced Markov chain $\{Y_t\}$ associated with the random variable $N_{n,k}$ is of the order $(l_n + 1)k \times (l_n + 1)k$. For more details, see Fu and Koutras (1994).

Koutras and Alexandrou (1995) defined $C_x = \{c_{x0}, c_{x1}, \ldots, c_{x,k-1}\}$, where

$$c_{xi} = \{(x, i)\}, \text{ for all } x \in S(X), \text{ and } 0 \le i \le k - 1.$$

This, along with the Markov chain $\{Y_t = (N_{t,k}, E_t)\}_{t=1}^{n}$ defined on Ω, the random variable $N_{n,k}$ becomes Markov chain imbeddable variable of binomial type, with

$$A_t(x) = A_t = \begin{bmatrix} (\cdot, 0) & (\cdot, 1) & (\cdot, 2) & \ldots & (\cdot, k-1) \\ q_t & p_t & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q_t & 0 & 0 & \ldots & p_t \\ q_t & 0 & 0 & \ldots & 0 \end{bmatrix}_{k \times k},$$

and

$$B_t(x) = B_t = \begin{bmatrix} (\cdot, 0) & (\cdot, 1) & (\cdot, 2) & \ldots & (\cdot, k-1) \\ 0 & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 \\ p_t & 0 & 0 & \ldots & 0 \end{bmatrix}_{k \times k}.$$

The advantage of using the Koutras and Alexandrou model to calculate the exact distribution of $N_{t,k}$ rather than Fu and Koutras model is the

fact that the Koutras and Alexandrou model deal with vector recurrences involving multiplications of $k \times k$ matrices instead of multiplying matrices of order $(l_n + 1)k \times (l_n + 1)k$.

Recently, Wu (2013) introduced "the double finite Markov chain imbedding technique" as an extension of the finite Markov chain imbedding technique, which makes it more flexible and helps to overcome the difficulties in finding the transition matrices. The main idea of the double finite Markov chain imbedding technique is to use the finite Markov chain imbedding technique repeatedly to obtain the transition probabilities of the induced Markov chain. Wu argued that, since each row of the probability transition matrix sums to 1, then such rows can be considered as a probability distributions for some random variables. If the random variables are Markov chain imbeddable, then we could use the finite Markov chain imbedding technique again. Wu (2013) used the double finite Markov chain imbedding technique to solve the coupon collector's problem.

Chapter 3

Runs and patterns in sequences

Let $\mathcal{S}_N = (s_1, s_2, \ldots, s_N)$ be a sequence of length N generated by selecting N integers from $\{1, 2, \ldots, n\}$ with unequal selection probabilities $p(s_i = j) = p_j$ for $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, n$, where p_j is the probability of selecting the integer "j" from $\{1, 2, \ldots, n\}$. For example, $\mathcal{S}_N = (4, 3, 5, 5, 2, 3, 1)$ is a sequence of length 7 generated from $\{1, 2, 3, 4, 5\}$ with equal selection probabilities $p_j = \frac{1}{5}$ for all $j = 1, \ldots, 5$. Let $A(\mathcal{S}_N, \ell)$ denote the total number of adjacent increasing subsequences of length ℓ in $\mathcal{S}_N$ (see Section 3.1), let $B(\mathcal{S}_N, \ell)$ denote the total number of adjacent consecutive increasing subsequences of length ℓ in $\mathcal{S}_N$ (see Section 3.2), and let $L(\mathcal{S}_N)$ denote the length of the longest increasing subsequence in $\mathcal{S}_N$ (see Section 3.3).

3.1 Adjacent increasing subsequences of length ℓ

For a given $\mathcal{S}_N = (s_1, s_2, \ldots, s_N)$, and $i = 1, \ldots, N - \ell + 1$. $(s_i, \ldots, s_{i+\ell-1})$ is said to be an adjacent increasing subsequence of length ℓ if $s_i < s_{i+1} < \cdots < s_{i+\ell-1}$. In this section we study the exact distribution of $A(\mathcal{S}_N, \ell)$ using the finite Markov chain imbedding technique.

If we are interested in finding the exact distribution of $A(\mathcal{S}_N, \ell)$ for $\ell \geq 2$ using the finite Markov chain imbedding technique, first consider the following state space:

$$\Omega = \{(a, r, j) : a = 0, 1, \ldots, m_{(N,n)}, r = 1, 2, \ldots, \ell - 1, \text{ and } r \leq j \leq n\},$$

where $m_{(N,n)}$ is the maximum number of overlapping adjacent increasing subsequences of length ℓ that can occur in $\mathcal{S}_N$ given by:

$$m_{(N,n)} = \begin{cases} N - \ell + 1 & n \geq N, \\ \lfloor N/n \rfloor (n - \ell + 1) + \max(0, N - n\lfloor N/n \rfloor - \ell + 1) & n < N. \end{cases}$$

Define a finite homogeneous Markov chain $\{Y_t\}_{t=1}^N$ on Ω as follows:

$$Y_t = (A(\mathcal{S}_t, \ell), R_t, J_t), \text{ for } 1 \leq t \leq N, \tag{3.1}$$

where $A(\mathcal{S}_t, \ell)$ is the total number of overlapping adjacent increasing subsequences of length ℓ that occurred within the first t elements $s_1, s_2, \ldots, s_t$, R_t is the length of the adjacent increasing subsequence ending at s_t, and J_t is the last element we observed (s_t).

24

For example, given $N = 8$, $n = 4$, $\ell = 3$, let $\mathcal{S}_8 = (3, 2, 4, 3, 4, 1, 2, 3)$, the realization of the imbedded Markov chain $\{Y_t\}_{t=1}^8$ with respect to this $\mathcal{S}_8$ is: $\{Y_1 = (0, 1, 3), Y_2 = (0, 1, 2), Y_3 = (0, 2, 4), Y_4 = (0, 1, 3), Y_5 = (0, 2, 4), Y_6 = (0, 1, 1), Y_7 = (0, 2, 2), Y_8 = (1, 3, 3)\}$. Note that, for a given sequence $\mathcal{S}_N$, the realization of $\{Y_t\}_{t=1}^N$ is always unique.

Proposition 3.1. *For a statement $\mathcal{S}$, let $\mathbf{1}\{\mathcal{S}\} = 1$ if $\mathcal{S}$ is true, and $\mathbf{1}\{\mathcal{S}\} = 0$ otherwise. For $t = 1, \ldots, N$, the transition probabilities $P_{(a',r',j'),(a,r,j)}$ are given by:*

$$P_{(a',r',j'),(a,r,j)} = \Pr(Y_t = (a, r, j) \mid Y_{t-1} = (a', r', j'))$$

$$= \begin{cases} p_j & \text{if } a = a' + \mathbf{1}\{r' = \ell - 1\}, j > j', \text{ and} \\ & \quad r = (r' + 1)\mathbf{1}\{r' < \ell - 1\} + (\ell - 1)\mathbf{1}\{r' = \ell - 1\} \\ p_j & \text{if } a = a' < m_{(N,n)}, j \leq j', \text{ and } r = 1, \\ 1 & \text{if } a = a' = m_{(N,n)}, j = j', \text{ and } r = r', \\ 0 & \text{otherwise.} \end{cases}$$

$$(3.2)$$

Proof: First note that, for all $j \leq j'$, the only transitions possible are $(a', r', j') \rightarrow (a', 1, j)$ and each of these happen with probability p_j. For all $j > j'$, we need to consider the value of r'. If $j > j'$ and $r' = \ell - 1$, then the number of adjacent increasing sequences increases by one and $r' \rightarrow r = \ell - 1$; otherwise, $r' \rightarrow r = r' + 1$ and $a' \rightarrow a = a'$. Each of these cases happen with with probability p_j. If $a' = m_{(N,n)}$, then we are in the absorbing state and we stay there with probability 1. $\qquad \square$

From Proposition 3.1, we have

$$\Pr(A(\mathcal{S}_N, \ell) = k) = \zeta_1 M^{N-1} U^\mathsf{T}(k), \qquad (3.3)$$

where ζ_1 is the probability vector of the first observation, M is the induced transition probability matrix, and $U(k)$ is a row vector with a 1 in the positions corresponding to the states $(k, \cdot, \cdot)$. In general, let

$$\mathscr{L}(A(\mathcal{S}_N, \ell)) = \zeta_1 M^{N-1} U^{*\mathsf{T}}, \tag{3.4}$$

where

$$\mathscr{L}(A(\mathcal{S}_N, \ell)) = [\Pr(A(\mathcal{S}_N, \ell) = 0), \ldots, \Pr(A(\mathcal{S}_N, \ell) = m_{(N,n)})],$$

and

$$U^* = \begin{bmatrix} U(0) \\ U(1) \\ \vdots \\ U(m_{(N,n)}) \end{bmatrix}.$$

Example 3.1. *Let $N = 4$, $n = 4$, and $\ell = 2$, with selection probabilities $p_1 = 0.20$, $p_2 = 0.30$, $p_3 = 0.35$, and $p_4 = 0.15$. Then the state space Ω is given by:*

$$\Omega = \{Y = (a, r, j); a = 0, 1, 2, 3, r = 1 \text{ and } 1 \leq j \leq 4\}$$

$$= \{(0,1,1), (0,1,2), (0,1,3), (0,1,4), (1,1,1), (1,1,2), (1,1,3), (1,1,4),$$

$$(2,1,1), (2,1,2), (2,1,3), (2,1,4), (3,1,1), (3,1,2), (3,1,3), (3,1,4)\},$$

and the transition matrix is given by

$$M = \begin{bmatrix} A & B & & \\ & A & B & \\ & & A & B \\ & & & I_4 \end{bmatrix},$$

26

where, I_k denotes the identity matrix of dimension $k \times k$, and

$$A = \begin{bmatrix} .20 & 0 & 0 & 0 \\ .20 & .30 & 0 & 0 \\ .20 & .30 & .35 & 0 \\ .20 & .30 & .35 & .15 \end{bmatrix}, \quad and \ B = \begin{bmatrix} 0 & .30 & .35 & .15 \\ 0 & 0 & .35 & .15 \\ 0 & 0 & 0 & .15 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For $\zeta_1 = (.20, .30, .35, .15, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$, and

$$U^* = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

we have

$$\mathscr{L}(A(\mathcal{S}_4, 2))^\mathsf{T} = \begin{bmatrix} \Pr(A(\mathcal{S}_4, 2) = 0) \\ \Pr(A(\mathcal{S}_4, 2) = 1) \\ \Pr(A(\mathcal{S}_4, 2) = 2) \\ \Pr(A(\mathcal{S}_4, 2) = 3) \end{bmatrix} = \begin{bmatrix} 0.1532563 \\ 0.6091375 \\ 0.2344562 \\ 0.0031500 \end{bmatrix}.$$

Example 3.2. *Let $N = 4$, $n = 4$ and $\ell = 3$, with equal selection probabilities $\frac{1}{4}$. Then the state space Ω is given by:*

$\Omega = \{Y = (a, r, j); u = 0, 1, 2, r = 1, 2, \ and \ r \leq j \leq 4\}$

$= \{(0, 1, 1), (0, 1, 2), (0, 1, 3), (0, 1, 4), (0, 2, 2), (0, 2, 3), (0, 2, 4), (1, 1, 1),$

$(1, 1, 2), (1, 1, 3), (1, 1, 4), (1, 2, 2), (1, 2, 3), (1, 2, 4), (2, 1, 1), (2, 1, 2),$

$(2, 1, 3), (2, 1, 4), (2, 2, 2), (2, 2, 3), (2, 2, 4)\},$

and the transition matrix is given by

$$M = \begin{bmatrix} A & B & \\ & A & B \\ & & I_7 \end{bmatrix},$$

27

where

$$A = \begin{bmatrix} .25 & 0 & 0 & 0 & .25 & .25 & .25 \\ .25 & .25 & 0 & 0 & 0 & .25 & .25 \\ .25 & .25 & .25 & 0 & 0 & 0 & .25 \\ .25 & .25 & .25 & .25 & 0 & 0 & 0 \\ .25 & .25 & 0 & 0 & 0 & 0 & 0 \\ .25 & .25 & .25 & 0 & 0 & 0 & 0 \\ .25 & .25 & .25 & .25 & 0 & 0 & 0 \end{bmatrix},$$

and

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & .25 & .25 \\ 0 & 0 & 0 & 0 & 0 & 0 & .25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For

$$\zeta_1 = (.25, .25, .25, .25, 0, 0, 0, 0, 0, 0, 0, 0, \ldots, 0, 0, 0, 0, 0, 0, 0, 0),$$

and

$$U^* = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

we have

$$\mathscr{L}(A(\mathcal{S}_4, 3))^{\mathsf{T}} = \begin{bmatrix} \Pr(A(\mathcal{S}_4, 3) = 0) \\ \Pr(A(\mathcal{S}_4, 3) = 1) \\ \Pr(A(\mathcal{S}_4, 3) = 2) \end{bmatrix} = \begin{bmatrix} 0.87890625 \\ 0.11718750 \\ 0.00390625 \end{bmatrix}.$$

Proposition 3.1 can be generalized for Markov dependent trials. If we are studying a sequence $\mathcal{S}_N$ generated by selecting N integers from

$\{1, 2, \ldots, n\}$ where the selection probabilities are Markov dependent with probability transition matrix P and $\zeta_1 = (\boldsymbol{\pi} \quad \mathbf{0}_{1 \times (N-n)})$, where $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_n)$ is the stationary distribution for the Markov dependent trials (i.e, $\boldsymbol{\pi} = \boldsymbol{\pi} P$), and $\mathbf{0}_{1 \times (N-n)}$ is a zero row vector, than the transition probabilities $P_{(a',r',j'),(a,r,j)}$ are given by the following proposition.

Proposition 3.2. *For $1 \leq t \leq N$, the transition probabilities $P_{(a',r',j'),(a,r,j)}$ are given by:*

$$P_{(a',r',j'),(a,r,j)} = \Pr(Y_t = (a,r,j) \mid Y_{t-1} = (a',r',j'))$$

$$= \begin{cases} \Pr(s_t = j \mid s_{t-1} = j') & \text{if } a = a' + \mathbf{1}\{r' = \ell - 1\}, j > j', \\ & \text{and } r = (r'+1)\mathbf{1}\{r' < \ell - 1\} + \\ & (\ell - 1)\mathbf{1}\{r' = \ell - 1\}, \\ \Pr(s_t = j \mid s_{t-1} = j') & \text{if } a = a' < m_{(N,n)}, j \leq j', \text{ and } r = 1, \\ 1 & \text{if } a = a' = m_{(N,n)}, j = j', \text{ and } r = r', \\ 0 & \text{otherwise.} \end{cases}$$

(3.5)

Proof: Identical to Proposition 3.1. $\qquad\square$

Example 3.3. *Let $N = 5$, $n = 3$, and $\ell = 2$, and let the selection from $\{1, 2, 3\}$ be governed by the probability transition matrix*

$$P = \begin{bmatrix} .35 & .25 & .40 \\ .25 & .15 & .60 \\ .25 & .25 & .50 \end{bmatrix}.$$

Then the state space Ω is given by:

$$\Omega = \{Y = (a, r, j); a = 0, 1, 2, 3, r = 1, \text{ and } 1 \leq j \leq 3\}$$

$$= \{(0,1,1), (0,1,2), (0,1,3), (1,1,1), (1,1,2), (1,1,3), (2,1,1),$$

$$(2,1,2), (2,1,3), (3,1,1), (3,1,2), (3,1,3)\},$$

and the transition matrix is given by

$$M = \begin{bmatrix} A & B & & \\ & A & B & \\ & & A & B \\ & & & I_3 \end{bmatrix},$$

where

$$A = \begin{bmatrix} .35 & 0 & 0 \\ .25 & .15 & 0 \\ .25 & .25 & .50 \end{bmatrix}, \ and \ B = \begin{bmatrix} 0 & .25 & .40 \\ 0 & 0 & .60 \\ 0 & 0 & 0 \end{bmatrix}.$$

For

$$\zeta_1 = (.2\bar{7}, .22\bar{72}, .4\bar{9}, 0, 0, 0, 0, 0, 0, 0, 0, 0),$$

and

$$U^* = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix},$$

we have

$$\mathscr{L}(A(\mathcal{S}_5, 2))^{\mathsf{T}} = \begin{bmatrix} \Pr(A(\mathcal{S}_5, 2) = 0) \\ \Pr(A(\mathcal{S}_5, 2) = 1) \\ \Pr(A(\mathcal{S}_5, 2) = 2) \\ \Pr(A(\mathcal{S}_5, 2) = 3) \end{bmatrix} = \begin{bmatrix} 0.1220092 \\ 0.5132102 \\ 0.3398753 \\ 0.0249053 \end{bmatrix}.$$

3.2 Adjacent consecutive increasing subsequences of length ℓ

An adjacent consecutive increasing subsequence of length ℓ occurs in $\mathcal{S}_N = (s_1, s_2, \ldots, s_N)$ if there exists $(s_i, \ldots, s_{i+\ell-1})$ such that $s_k = s_{k-1} + 1$

for each $k = i, \ldots, i+\ell-1$. In this section we study the exact distribution of $B(\mathcal{S}_N, \ell)$ using the finite Markov chain imbedding technique.

If we are interested in finding the exact distribution of $B(\mathcal{S}_N, \ell)$ for $\ell \geq 2$ using the finite Markov chain imbedding technique, first consider the following state space:

$$\Omega = \{(b, r, j) : b = 0, 1, \ldots, e_{(N,n)}, r = 1, 2, \ldots, \ell-1, \text{ and } r \leq j \leq n\}$$

where $e_{(N,n)}$ is the maximum number of overlapping adjacent consecutive increasing subsequences of length ℓ that can occur in $\mathcal{S}_N$ given by:

$$e_{(N,n)} = \begin{cases} N - \ell + 1 & n \geq N, \\ \lfloor N/n \rfloor (n - \ell + 1) + \max(0, N - n\lfloor N/n \rfloor - \ell + 1) & n < N. \end{cases}$$

Define a finite homogeneous Markov chain $\{Y_t\}_{t=1}^N$ on Ω as follows:

$$Y_t = (B(\mathcal{S}_t, \ell), R_t, J_t), \text{ for } 1 \leq t \leq N, \tag{3.6}$$

where $B(\mathcal{S}_t, \ell)$ is the total number of overlapping adjacent consecutive increasing subsequences of length ℓ that occurred within the first t elements $s_1, s_2, \ldots, s_t$, R_t is the length of the adjacent increasing subsequence associated with s_t, and J_t is the last element we observe (s_t). For example, given $N = 8$, $n = 5$, and $\ell = 3$ let $\mathcal{S}_8 = (1, 3, 2, 4, 3, 4, 5, 1)$, the realization of the imbedded Markov chain $\{Y_t\}_{t=1}^8$ with respect to this $\mathcal{S}_8$ is: $\{Y_1 = (0, 1, 1), Y_2 = (0, 1, 3), Y_3 = (0, 1, 2), Y_4 = (0, 1, 4), Y_5 = (0, 1, 3), Y_6 = (0, 2, 4), Y_7 = (1, 3, 5), Y_8 = (1, 1, 1)\}$. Note that for a given sequence $\mathcal{S}_N$, the realization of $\{Y_t\}_{t=1}^N$ is always unique.

31

Proposition 3.3. *For $1 \leq t \leq N$, the transition probabilities $P_{(b',r',j'),(b,r,j)}$*

are given by:

$$P_{(b',r',j'),(b,r,j)} = \Pr(Y_t = (b,r,j) \mid Y_{t-1} = (b',r',j'))$$

$$= \begin{cases} p_j & \text{if } b = b' + \mathbf{1}\{r' = \ell - 1\}, j = j' + 1, \text{ and} \\ & \quad r = (r'+1)\mathbf{1}\{r' < \ell - 1\} + (\ell - 1)\mathbf{1}\{r' = \ell - 1\}, \\ p_j & \text{if } b = b' < e_{(N,n)}, j \neq j' + 1, \text{ and } r = 1, \\ 1 & \text{if } b = b' = e_{(N,n)}, j = j', \text{ and } r = r', \\ 0 & \text{otherwise.} \end{cases}$$

$$(3.7)$$

Proof: First note that, for all $j \neq j' + 1$, the only transitions possible are

$(b', r', j') \rightarrow (b', 1, j)$ and each of these happen with probability p_j. For

$j = j' + 1$, we need to consider the value of r'. If $j = j' + 1$ and $r' = \ell - 1$,

then the number of adjacent consecutive increasing sequences increases by

one and $r' \rightarrow r = \ell - 1$; otherwise, $r' \rightarrow r = r' + 1$ and $b' \rightarrow b = b'$. Each

of these cases happen with with probability p_j. If $b' = e_{(N,n)}$, then we are

in the absorbing state and we stay there with probability 1. $\square$

Form Proposition 3.3, we have

$$\Pr(B(\mathcal{S}_N, \ell) = k) = \zeta_1 M^{N-1} U^\mathsf{T}(k), \tag{3.8}$$

where ζ_1 is the probability vector of the first observation, M is transition

probability matrix, and $U(k)$ is a row vector with a 1 in the positions

corresponding to the states $(k, \cdot, \cdot)$. In general let

$$\mathscr{L}(B(\mathcal{S}_N, \ell)) = \zeta_1 M^{N-1} U^{*\mathsf{T}}, \tag{3.9}$$

where

$$\mathscr{L}(B(\mathcal{S}_N, \ell)) = [\Pr(B(\mathcal{S}_N, \ell) = 0), \dots, \Pr(B(\mathcal{S}_N, \ell) = e_{(N,n)})],$$

and

$$U^* = \begin{bmatrix} U(0) \\ U(1) \\ \vdots \\ U(e_{(N,n)}) \end{bmatrix}.$$

Example 3.4. *Let $N = 5$, $n = 4$, and $\ell = 2$, with selection probabilities $p_1 = 0.20$, $p_2 = 0.30$, $p_3 = 0.35$, and $p_4 = 0.15$. Then the state space Ω is given by:*

$$\Omega = \{Y = (b, r, j); b = 0, 1, 2, 3, r = 1, \text{ and } 1 \leq j \leq 4\}$$

$$= \{(0, 1, 1), (0, 1, 2), (0, 1, 3), (0, 1, 4), (1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 1, 4),$$

$$(2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 1, 4), (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 1, 4)\},$$

and the transition matrix is given by

$$M = \begin{bmatrix} A & B & & \\ & A & B & \\ & & A & B \\ & & & I_4 \end{bmatrix},$$

where

$$A = \begin{bmatrix} .20 & 0 & .35 & .15 \\ .20 & .30 & 0 & .15 \\ .20 & .30 & .35 & 0 \\ .20 & .30 & .35 & .15 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 0 & .30 & 0 & 0 \\ 0 & 0 & .35 & 0 \\ 0 & 0 & 0 & .15 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

For

$$\zeta_1 = (.20, .30, .35, .15, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),$$

33

and

$$U^* = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

we have

$$\mathscr{L}(B(\mathcal{S}_5, 2))^{\mathsf{T}} = \begin{bmatrix} \Pr(B(\mathcal{S}_5, 2) = 0) \\ \Pr(B(\mathcal{S}_5, 2) = 1) \\ \Pr(B(\mathcal{S}_5, 2) = 2) \\ \Pr(B(\mathcal{S}_5, 2) = 3) \end{bmatrix} = \begin{bmatrix} 0.35988250 \\ 0.43252120 \\ 0.18531000 \\ 0.02228625 \end{bmatrix}.$$

Example 3.5. *Let $N = 5$, $n = 4$ and $\ell = 3$, with equal selection probability $\frac{1}{4}$. Then the state space Ω is given by:*

$$\Omega = \{Y = (b, r, j) : b = 0, 1, 2, r = 1, 2, \text{ and } r \le j \le 4\}$$

$$= \{(0, 1, 1), (0, 1, 2), (0, 1, 3), (0, 1, 4), (0, 2, 2), (0, 2, 3), (0, 2, 4), (1, 1, 1),$$

$$(1, 1, 2), (1, 1, 3), (1, 1, 4), (1, 2, 2), (1, 2, 3), (1, 2, 4), (2, 1, 1), (2, 1, 2),$$

$$(2, 1, 3), (2, 1, 4), (2, 2, 2), (2, 2, 3), (2, 2, 4)\},$$

and the transition matrix is given by

$$M = \begin{bmatrix} A & B & \\ & A & B \\ & & I_7 \end{bmatrix},$$

where

$$A = \begin{bmatrix} .25 & 0 & .25 & .25 & .25 & 0 & 0 \\ .25 & .25 & 0 & .25 & 0 & .25 & 0 \\ .25 & .25 & .25 & 0 & 0 & 0 & .25 \\ .25 & .25 & .25 & .25 & 0 & 0 & 0 \\ .25 & .25 & 0 & .25 & 0 & 0 & 0 \\ .25 & .25 & .25 & 0 & 0 & 0 & 0 \\ .25 & .25 & .25 & .25 & 0 & 0 & 0 \end{bmatrix},$$

34

and

$$
B = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .25 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & .25 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
$$

For

$$
\zeta_1 = (.25, .25, .25, .25, 0, 0, 0, 0, 0, 0, 0, \ldots, 0, 0, 0, 0, 0, 0, 0),
$$

and

$$
U^* = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1
\end{bmatrix},
$$

we have

$$
\mathscr{L}(B(\mathcal{S}_5, 3))^{\mathsf{T}} = \begin{bmatrix}
\Pr(B(\mathcal{S}_5, 3) = 0) \\
\Pr(B(\mathcal{S}_5, 3) = 1) \\
\Pr(B(\mathcal{S}_5, 3) = 2)
\end{bmatrix} = \begin{bmatrix}
0.9140625 \\
0.0781250 \\
0.0078125
\end{bmatrix}.
$$

Proposition 3.3 can be generalized for Markov dependent trials. Again, if we are studying a sequence $\mathcal{S}_N$ generated by selecting N integers from $\{1, 2, \ldots, n\}$ where the selection probabilities are Markov dependent with probability transition matrix P and $\zeta_1 = (\boldsymbol{\pi} \quad \mathbf{0}_{1 \times (N-n)})$, where $\boldsymbol{\pi} = (\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_n)$ is the stationary distribution for the Markov dependent trials (i.e, $\boldsymbol{\pi} = \boldsymbol{\pi} P$), and $\mathbf{0}_{1 \times (N-n)}$ is a zero row vector, then the transition probabilities $P_{(b', r', j'), (b, r, j)}$ are given by the following proposition.

Proposition 3.4. *For $1 \le t \le N$, the transition probabilities $P_{(b',r',j'),(b,r,j)}$*

are given by:

$$P_{(b',r',j'),(b,r,j)} = \Pr(Y_t = (b,r,j) \mid Y_{t-1} = (b',r',j'))$$

$$= \begin{cases} \Pr(s_t = j \mid s_{t-1} = j') & \text{if } b = b' + \mathbf{1}\{r' = \ell - 1\}, j = j' + 1, \text{ and} \\ & r = (r'+1)\mathbf{1}\{r' < \ell - 1\} + \\ & (\ell - 1)\mathbf{1}\{r' = \ell - 1\}, \\ \Pr(s_t = j \mid s_{t-1} = j') & \text{if } b = b' < e_{(N,n)}, j \ne j' + 1, \text{ and } r = 1, \\ 1 & \text{if } b = b' = e_{(N,n)}, j = j', \text{ and } r = r', \\ 0 & \text{otherwise.} \end{cases}$$

$$(3.10)$$

Proof: Identical to Proposition 3.3. $\qquad\qquad\square$

Example 3.6. *Let $N = 5$, $n = 4$, and $\ell = 2$, and let the selection from $\{1, 2, 3, 4\}$ be governed by the probability transition matrix*

$$P = \begin{bmatrix} .25 & .25 & .10 & .40 \\ .20 & .30 & .35 & .15 \\ .30 & .15 & .10 & .45 \\ .20 & .10 & .30 & .40 \end{bmatrix}.$$

Then the state space Ω is given by:

$$\Omega = \{Y = (b,r,j); b = 0, 1, 2, 3, r = 1 \text{ and } 1 \le j \le 4\}$$

$$= \{(0,1,1), (0,1,2), (0,1,3), (0,1,4), (1,1,1), (1,1,2), (1,1,3),$$

$$(1,1,4), (2,1,1), (2,1,2), (2,1,3), (2,1,4), (3,1,1), (3,1,2),$$

$$(3,1,3), (3,1,4)\},$$

and the transition matrix is given by

$$
M = \begin{bmatrix} A & B & & \\ & A & B & \\ & & A & B \\ & & & I_4 \end{bmatrix},
$$

where

$$
A = \begin{bmatrix} .25 & 0 & .10 & .40 \\ .20 & .30 & 0 & .15 \\ .30 & .15 & .10 & 0 \\ .20 & .10 & .30 & .40 \end{bmatrix}, \ and \ B = \begin{bmatrix} 0 & .25 & 0 & 0 \\ 0 & 0 & .35 & 0 \\ 0 & 0 & 0 & .45 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
$$

For

$$
\zeta_1 = (.23, .18, .22, .37, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0),
$$

and

$$
U^* = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},
$$

we have

$$
\mathscr{L}(B(\mathcal{S}_5, 2))^{\mathsf{T}} = \begin{bmatrix} \Pr(B(\mathcal{S}_5, 2) = 0) \\ \Pr(B(\mathcal{S}_5, 2) = 1) \\ \Pr(B(\mathcal{S}_5, 2) = 2) \\ \Pr(B(\mathcal{S}_5, 2) = 3) \end{bmatrix} = \begin{bmatrix} 0.37627650 \\ 0.40331770 \\ 0.18192950 \\ 0.03847639 \end{bmatrix}.
$$

As we can see from the previous examples, our homogeneous finite Markov chain is of binomial type, then we can use the results in Koutras and Alexandrou (1995) to obtain the generating functions and means for $A(\mathcal{S}_N, \ell)$ and $B(\mathcal{S}_N, \ell)$.

3.3 The length of the longest increasing subsequence

The length of the longest increasing subsequence in a sequence of n distinct elements generated from a linearly ordered set is determined via algorithms. Fredman (1975) described an algorithm which determine the length of the longest increasing subsequence. Fredman shows that for any algorithm to provide sufficient information to compute the length of the longest increasing subsequence in a sequence of n distinct elements it must perform at least $[n \log n - n \log \log n + O(n)]$ comparisons in its worst case. There are many dynamic programming algorithms to compute the length of the longest increasing subsequence (see, for example, Fredman, 1975; Gonnet and Baeza-Yates, 1991; Sun and Woodruff, 2007). Even though, these algorithms apparently are efficient at computing the length of the longest increasing subsequence in a sequence of n distinct elements from a linearly ordered set, they do not necessarily give insight into the distribution of the length of the longest increasing subsequences in such sequences. The upper bound of the exact distribution of the length of the longest increasing subsequence of a sequence of n distinct elements has been studied extensively (see, for example, McDiarmid, 1998).

Let $\mathcal{S}_N = (s_1, s_2, \ldots, s_N)$ be a sequence of length N generated by selecting N integers from $\{1, 2, \ldots, n\}$ with equal selection probabilities, and let $L(\mathcal{S}_N)$ denote the length of the longest increasing subsequence in $\mathcal{S}_N$. The exact distribution of the length of the longest increasing

subsequence in sequences can be calculated as the exact distribution of the length of the longest increasing subsequence in random permutation of arbitrary multi-sets. For a given $\mathcal{S}_N$ we can calculate the exact distribution of $L(\mathcal{S}_N)$ as the exact distribution of $L_N([\boldsymbol{d}_n], \pi)$ (for more details, see Section 4.1). Generally, the exact distribution of the length of the longest increasing subsequence in the set of all sequences of length N generated from $\{1, 2, \ldots, n\}$ can be calculated as the exact distribution of $L_N(\prod(n))$ (for more details, see Section 4.2).

3.4 Alternative methods

A variety of methods have been proposed for studying the exact distribution for runs and patterns in sequences, none of the available methods is superior. Fu and Lou (2003) used the finite Markov chain imbedding technique along with the forward and backward principle to calculate the exact distribution of the number of compound patterns in sequences of multi-state trials. To compare between the Fu and Lou (2003) approach and our approach, consider the case of $A(\mathcal{S}_N, \ell)$, for $N = 4$, $n = 4$, and $\ell = 3$. Then the state space Ω for the Fu and Lou (2003) approach is given by:

$$
\begin{aligned}
\Omega = \{ & \emptyset, (0, 1), (0, 2), (0, 3), (0, 4), (0, 12), (0, 13), (0, 23), (1, 1), (1, 2), (1, 3), \\
& (1, 4), (1, 12), (1, 23), (1, 13), (1, 123), (1, 124), (1, 134), (1, 234), (2, 1), \\
& (2, 2), (2, 3), (2, 4), (2, 12), (2, 23), (2, 13), (2, 123), (2, 124), (2, 134), \\
& (2, 234) \},
\end{aligned}
$$

while, the state space Ω for our approach is given by

$$
\begin{aligned}
\Omega = \{ & (0,1,1), (0,1,2), (0,1,3), (0,1,4), (0,2,2), (0,2,3), (0,2,4), (1,1,1), \\
& (1,1,2), (1,1,3), (1,1,4), (1,2,2), (1,2,3), (1,2,4), (2,1,1), (2,1,2), \\
& (2,1,3), (2,1,4), (2,2,2), (2,2,3), (2,2,4) \},
\end{aligned}
$$

this state space is smaller than Fu and Lou (2003) state space. Also, in our approach, $A(\mathcal{S}_N, \ell)$ is Markov chain imbeddable variable of binomial type.

Also, Fu and Lou (2003) used the finite Markov chain imbedding technique for the compound patterns to find the probability $\Pr(L(\mathcal{S}_N) < v)$. Let $\Lambda(v)$ be a compound pattern given by $\Lambda(v) = \bigcup_{i=1}^{j} \Lambda_i$, where Λ_i's are all possible distinct increasing v-sequences that could be generated from $\{1, 2, \ldots, n\}$. Fu and Lou (2003) used the dual relationship between $L(\mathcal{S}_N)$ and the waiting time for the first occurrence of an increasing subsequence of length v denoted by $W(\Lambda(v))$,

$$
\{L(\mathcal{S}_N) < v\} \text{ if and only if } \{W(\Lambda(v)) > N\},
$$

to calculate $\Pr(L(\mathcal{S}_N) < v)$ as follows

$$
\Pr(L(\mathcal{S}_N) < v) = \Pr(W(\Lambda(v)) > N).
$$

Nicodeme et al. (2002) proposed using deterministic finite automata (DFAs) in order to study the exact distribution for run-related statistics. Regular expressions can be represented by minimal state deterministic

40

finite automata or finite state machines, and there exists algorithms for performing the conversion, which allow us to construct a finite Markov chain imbedding to study the distributional characteristics of these patterns in iid sequences of multi-state trials (see, for example, Hopcroft et al., 2001; Crochemore and Stefanov, 2003; Stefanov et al., 2007). A deterministic finite automaton is a 5-tuple $\mathcal{D} = (\mathcal{S}, \mathcal{A}, \delta, s_0, \mathcal{F})$, where $\mathcal{S}$ denotes a finite set of states, $\mathcal{A}$ denotes a finite alphabet, δ denotes a transition function (i.e. $\delta : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$), s_0 is an initial state, and $\mathcal{F}$ denotes a set of accepting states. A deterministic finite automaton $\mathcal{D}$ is conveniently represented as a directed graph or state transition diagram, where the states in $\mathcal{S}$ are the nodes of the graph, with the accepting states $\mathcal{F}$ depicted by double circles and, for every state $s \in \mathcal{S}$ and every character $\alpha \in \mathcal{A}$, there is a direct edge from s to $\delta(s, \alpha)$ with α as part of it's label. For example, for $n = 3$ the state transition diagram for the waiting time to the first increasing 3-sequences shown in Figure 3.1. Then by constructing a deterministic finite automaton $\mathcal{D}$ with transition function δ, the Markov chain induced by $\mathcal{D}$ gives the distribution of the waiting time $W(\Lambda)$ for a regular expression.
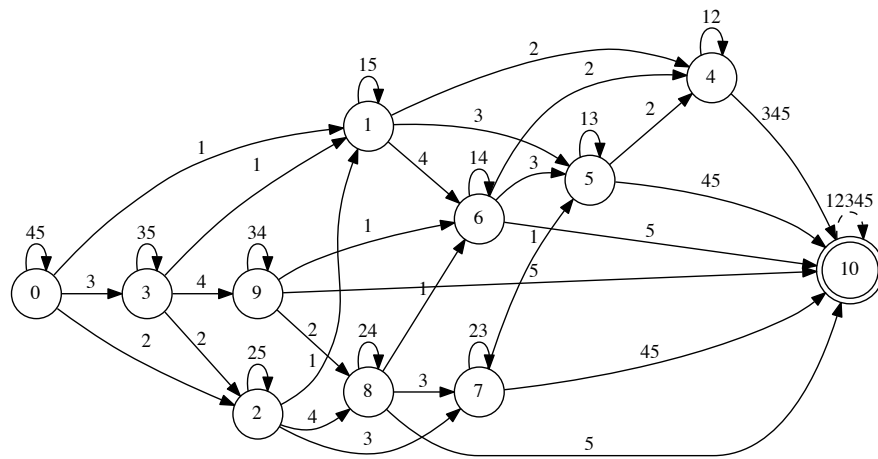
Figure 3.1: The state transition diagram for $n = 3$ and $v = 3$

Chapter 4

The length of the longest increasing subsequence in random permutations of arbitrary multi-sets

In this chapter we study the exact distribution of the length of the longest increasing subsequence in random permutations of arbitrary multi-sets using the finite Markov chain imbedding technique. Numerical examples are provided to aid in understanding. First, we study the conditional case. Given $[\boldsymbol{d}_n]$, and $\pi \in \prod[\boldsymbol{d}_n]$, the exact distribution of the length of the longest increasing subsequence in π is given in Section 4.1. Then, we study the unconditional case. In Section 4.2 we introduce the exact distribution of the length of the longest increasing subsequence for the set of all permutations of length N generated from $\{1, 2, \ldots, n\}$.

4.1 The conditional case

Recall from Chapter 1, that $L_N([\boldsymbol{d}_n], \pi)$ denotes the length of the longest increasing subsequence in a random permutation $\pi \in \prod[\boldsymbol{d}_n]$. Fu and Hsieh (2015) obtained the exact distribution of $L_n([\mathbf{1}_n], \pi)$ using the finite Markov chain imbedding technique. Expanding their result, we obtain the exact distribution of the length of the longest increasing subsequence in random permutations of arbitrary multi-sets $L_N([\boldsymbol{d}_n], \pi)$. We will begin with some notation.

Definition 4.1. *For a given $[\boldsymbol{d}_n]$, and $\pi(t)$, we define the vector $[\boldsymbol{a}_b(t)] = [a_1(t), \ldots, a_b(t)]$ where for each $j = 1, \ldots, b$ and each $1 \leq b \leq k(t)$, $a_j(t)$ is the position of the last integer in the leftmost increasing subsequence of length j in $\pi(t)$.*

Note that, $a_1(t) = 1 < a_2(t) < \cdots < a_b(t) \leq t$, and there may be more than one $\pi(t)$ with the same $[\boldsymbol{a}_b(t)]$.

Definition 4.2. *For a given $[\boldsymbol{d}_n]$, and $\pi(t)$ with $[\boldsymbol{a}_b(t)]$, we define $[\boldsymbol{x}_z(t)] = [x_1(t), \ldots, x_z(t)]$ where for $i = 1, \ldots, z$ and $1 \leq z \leq k(t) - 1$, $x_i(t)$ is the position of the last integer of the leftmost increasing subsequence of length i in $\pi(t)$, which ends with an integer less than the integer "$k(t)$".*

Note that, when no such increasing subsequences exist, $[\boldsymbol{x}_z(t)] = 0$. It is clear that $1 \leq x_z(t) \leq t$. To illustrate these definitions, consider the following example: let $[\boldsymbol{d}_8] = [1, 1, 1, 1, 1, 1, 1, 3]$ and $\pi(10) = (8634851278)$; then (8) is the leftmost increasing subsequence of length 1 and $a_1(10) = 1$,

(34) is the leftmost increasing subsequence of length 2 and $a_2(10) = 4$, (348) is the leftmost increasing subsequence of length 3 and $a_3(10) = 5$, (3457) is the leftmost increasing subsequence of length 4 and $a_4(10) = 9$, and (34578) is the leftmost increasing subsequence of length 5 and $a_5(10) = 10$. Thus $[\boldsymbol{a}_5(10)] = [1, 4, 5, 9, 10]$. Moreover, (6) is the leftmost increasing subsequence of length 1 which ended with integer less than 8 and $x_1(10) = 2$, (34) is the leftmost increasing subsequence of length 2 which ended with integer less than 8 and $x_2(10) = 4$, (345) is the leftmost increasing subsequence of length 3 which ended with integer less than 8 and $x_3(10) = 6$, and (3457) is the leftmost increasing subsequence of length 4 which ended with integer less than 8 and $x_4(10) = 9$. Then $[\boldsymbol{x}_4(10)] = [2, 4, 6, 9]$.

Note that, the leftmost increasing subsequence of length i in $\pi(t)$ is the first increasing subsequence of length i in $\pi(t)$. If there are several increasing subsequences of length i ending with the same integer, then the leftmost increasing subsequence is the one that started with the integer furthest to the left in $\pi(t)$. For example, let $[\boldsymbol{d}_6] = [2, 2, 2, 3, 2, 1]$ and $\pi(12) = (421134562354)$ then (23456), (13456) and (13456) are increasing subsequences of length 5 ended with 6, but the leftmost increasing subsequence of length 5 is the one starting with 2.

Given $\pi(t-1)$, let $\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t, i)\rangle = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)])$ denote the mapping induced by inserting the $c(t)$-th copy of the integer "$k(t)$" into the i-th gap of the permutation $\pi(t-1)$, including the two end gaps. The following (very long) proposition specifies the relationship

45

between $([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)])$ and $([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)])$.

Proposition 4.1. *For given* $([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)])$, $1 \le b' \le t-1$ *and* $1 \le z' \le b'$.

(1) If $i = 1$, *and* $c(t) = 1$, *then*

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t, i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{b'}(t)]),$$

where $a_1(t) = 1$, $a_l(t) = a_l(t-1)+1$ *for* $l = 2, \ldots, b'$, *and* $x_k(t) = a_k(t-1)+1$ *for* $k = 1, \ldots, b'$.

(2) If $a_{j-1}(t-1) < i \le a_j(t-1)$ *for some* $j = 2, \ldots, b'$, *and* $c(t) = 1$, *then*

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t, i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{b'}(t)]),$$

where

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, j-1, \\ i & \text{for } l = j, \\ a_l(t-1)+1 & \text{for } l = j+1, \ldots, b'; \end{cases}$$

and

$$x_r(t) = \begin{cases} a_r(t-1) & \text{for } r = 1, \ldots, j-1, \\ a_r(t-1)+1 & \text{for } r = j, \ldots, b'. \end{cases}$$

(3) If $a_{b'}(t-1) < i \le t$, *and* $c(t) = 1$, *then*

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t, i)\rangle = ([\boldsymbol{a}_{b'+1}(t)], [\boldsymbol{x}_{b'}(t)]),$$

where

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, b', \\ i & \text{for } l = b'+1; \end{cases}$$

and

$$x_r(t) = a_r(t-1), \ \text{for } r = 1, \ldots, b'.$$

(4) *If* $i = 1$, *and* $c(t) > 1$, *then*

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t,i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)]),$$

where $a_1(t) = 1$, $a_l(t) = a_l(t-1) + 1$ *for* $l = 2, \ldots, b'$, *and* $x_r(t) = x_r(t-1) + 1$ *for* $r = 1, \ldots, z'$.

(5) *If* $a_{j-1}(t-1) < i \le a_j(t-1)$ *for* $j = 2, \ldots, b'$, $c(t) > 1$, *then*

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t,i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)]),$$

where

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, j-1, \\ i & \text{for } l = j, \text{ and } a_{j-1}(t-1) = x_{j-1}(t-1) \text{ or} \\ & \quad a_{j-1}(t-1) < x_{j-1}(t-1) < i, \\ a_l(t-1) + 1 & \text{for } l = j, a_{j-1}(t-1) < x_{j-1}(t-1), \\ & \quad x_{j-1}(t-1) \ge i, \\ a_l(t-1) + 1 & \text{for } l = j+1, \ldots, b'; \end{cases}$$

and, for $r = 1, \ldots, z'$, *let*

$$m = \begin{cases} 0 & \text{for } i \le x_1(t-1), \\ \max(d : 1 \le d \le z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \\ & \quad \le x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1); \end{cases}$$

and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } 1 \le r \le m, \\ x_r(t-1) + 1 & \text{for } m < r \le z'. \end{cases}$$

47

(6) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) > x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t,i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)]),$$

where

$$a_l(t) = a_l(t-1), \ \ for \ l = 1, \ldots, b',$$

and

$$x_r(t) = x_r(t-1), \ \ for \ r = 1, \ldots, z'.$$

(7) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) = x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t,i)\rangle = ([\boldsymbol{a}_{b'+1}(t)], [\boldsymbol{x}_{z'}(t)]),$$

where

$$a_l(t) = \begin{cases} a_l(t-1) & for \ l = 1, \ldots, b', \\ i & for \ l = b' + 1; \end{cases}$$

and

$$x_r(t) = x_r(t-1), \ \ for \ r = 1, \ldots, z'.$$

(8) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) < x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]), (t,i)\rangle = ([\boldsymbol{a}_{b}(t)], [\boldsymbol{x}_{z'}(t)]),$$

where $b = b'$ if $x_{z'}(t-1) \ge i$ or $b = b' + 1$ if $x_{z'}(t-1) < i$, and

$$a_l(t) = \begin{cases} a_l(t-1) & for \ l = 1, \ldots, b', \\ i & for \ l = b' + 1, \ \ and \ x_{z'}(t-1) < i; \end{cases}$$

48

and, for $r = 1, \ldots, z'$, let

$$
m = \begin{cases}
0 & \textit{for } i \leq x_1(t-1), \\
\max(d : 1 \leq d \leq z', x_d(t-1) < i) & \textit{for } x_1(t-1) < i \\
& \quad \leq x_{z'}(t-1), \\
i & \textit{for } i > x_{z'}(t-1);
\end{cases}
$$

and

$$
x_r(t) = \begin{cases}
x_r(t-1) & \textit{for } 1 \leq r \leq m, \\
x_r(t-1) + 1 & \textit{for } m < r \leq z'.
\end{cases}
$$

Proof: When $c(t) = 1$, then the integer "$k(t)$" is the largest integer among all the integers in $\pi(t-1)$. If $i = 1$ then $a_1(t) = 1$, and all the leftmost increasing subsequences of length greater than 1 are the same as in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as the leftmost increasing subsequences in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right, which proves part (1).

If $c(t) = 1$, and $a_{j-1}(t-1) < i \leq a_j(t-1)$, then all the leftmost increasing subsequences with length less than or equal to $j - 1$ are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. While, the leftmost increasing subsequence associated with $a_{j-1}(t-1)$, together with the integer $k(t)$, will form a leftmost increasing subsequence of length j with $a_j(t) = i$. All the leftmost increasing subsequences of length greater than j are the

49

same as in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as the leftmost increasing subsequences in $\pi(t-1)$, and the positions of the endpoints of any leftmost increasing subsequences of length less than j also stay the same as in $\pi(t-1)$, otherwise they will shift to the right, which proves part (2).

If $c(t) = 1$, and $a_{b'}(t-1) < i \leq t$, then all the leftmost increasing subsequences of length less than or equal to b' are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. While, the leftmost increasing subsequence associated with $a_{b'}(t-1)$, together with the integer $k(t)$, will form a leftmost increasing subsequence of length $b' + 1$ with $a_{b'+1}(t) = i$. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as the leftmost increasing subsequences in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same, which proves part (3).

When $c(t) > 1$, then the integer $k(t)$ is equal to at least one of the integers in $\pi(t-1)$. If $i = 1$, then $a_1(t) = 1$, and all the leftmost increasing subsequences of length greater than 1 are the same as in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right,

50

which proves part (4).

If $c(t) > 1$, and $a_{j-1}(t-1) < i \leq a_j(t-1)$, then all the leftmost increasing subsequences of length less than or equal to $j-1$ are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. If the leftmost increasing subsequence of length $j-1$ which ends with integer less than $k(t)$ ended before the insertion position i, then the leftmost increasing subsequence of length $j-1$ in $\pi(t-1)$, together with the integer $k(t)$, will form a leftmost increasing subsequence of length j with $a_j(t) = i$, otherwise $a_j(t) = a_j(t-1) + 1$. While, all the leftmost increasing subsequences of length greater than j are the same as in $\pi(t-1)$, but the positions of the endpoints of these increasing subsequences will shift to the right. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as the leftmost increasing subsequences in $\pi(t-1)$, and for any of these increasing subsequences which ends before the insertion position i the positions of the endpoints also stay the same, otherwise they will shift to the right, which proves part (5).

If $c(t) > 1$, $a_{b'}(t-1) < i \leq t$, and $a_{b'}(t-1) > x_{z'}(t-1)$, then all the leftmost increasing subsequences of length less than or equal to b' are the same in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same, which proves part (6).

If $c(t) > 1$, $a_{b'}(t-1) < i \le t$, and $a_{b'}(t-1) = x_{z'}(t-1)$, then all the leftmost increasing subsequences of length less than or equal to b' are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. Since $a_{b'}(t-1) = x_{z'}(t-1)$, the leftmost increasing subsequence of length b' ended with integer less than $k(t)$, then the leftmost increasing subsequence associated with $a_{b'}(t-1)$, together with the integer $k(t)$, will form a leftmost increasing subsequence of length $b' + 1$ with $a_{b'+1}(t) = i$. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same, which proves part (7).

If $c(t) > 1$, $a_{b'}(t-1) < i \le t$, and $a_{b'}(t-1) < x_{z'}(t-1)$, then all the leftmost increasing subsequences of length less than or equal to b' are the same as in $\pi(t-1)$, and the positions of the endpoints of these increasing subsequences also stay the same. Since $a_{b'}(t-1) < x_{z'}(t-1)$, we must have $z' = b'$, and if $x_{z'}(t-1) < i$ then the leftmost increasing subsequence associated with $x_{z'}(t-1)$, together with the integer $k(t)$, will form a leftmost increasing subsequence of length $b' + 1$ with $a_{b'+1}(t) = i$, otherwise the length of the longest increasing subsequence stays the same. For $[\boldsymbol{x}_{z'}(t)]$, the leftmost increasing subsequences which end with an integer less than $k(t)$ are the same as in $\pi(t-1)$, and the positions of the endpoints of any leftmost increasing subsequences of length less than or equal to i also stay the same, otherwise they will shift to the right, which proves part (8). □

Definition 4.3. *Let v denote the length of the longest increasing subsequence of interest. Then for a given N, v, and $t = 1, \ldots, N$, we define*

$$\underline{B}_t = max(1, v + t - N) \ and \ \bar{B}_t = min(k(t), v - 1).$$

Note that, for $1 \leq t \leq N - 1$, $\underline{B}_t \leq \bar{B}_t$, but for $t = N$, $\underline{B}_N > \bar{B}_N$.

Definition 4.4. *A state $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) \in \Omega_t$ is defined as a lower critical state if $b = v + t - N$ (i.e. $\underline{B}_t = max(1, v + t - N) = v + t - N$) and is defined as an upper critical state if $b = v - 1$ (i.e. $\bar{B}_t = min(k(t), v - 1) = v - 1$), otherwise it is referred as a regular state.*

The critical states are the only states that could be absorbed. If $b = \underline{B}_t$ then Y_t could be absorbed into $\emptyset$. But if $b = \bar{B}_t$ then Y_t could be absorbed into α. Now we are ready to defined the sequence of the state space $\{\Omega\}_{t=1}^N$ by

$$\Omega_t = \{([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) : b \geq \underline{B}_t, b \leq \bar{B}_t \text{ and } 1 \leq z \leq b\} \bigcup \{\emptyset, \alpha\},$$

where $\{\Omega_t\}_{t=1}^N$ is the sequence of the states spaces for the non-homogeneous Markov chain $\{Y_t\}_{t=1}^N$, where the state $\{\alpha\}$ represents an absorbing state such that if $Y_t \in \alpha$ then the length of the longest increasing subsequence is greater or equal to v, and the state $\{\emptyset\}$ represent an absorbing state such that if $Y_t \in \emptyset$ then the length of the longest increasing subsequence is less than v. Using $\underline{B}_t$ and $\bar{B}_t$ makes sure that the state spaces $\{\Omega_t\}_{t=1}^N$ are as small as possible. We always start with $\Omega_1 = \{\emptyset, ([1], 0), \alpha\}$, and since $\underline{B}_N > \bar{B}_N$, $\Omega_N = \{\emptyset, \alpha\}$. While, generating the other states depends

53

on $[\boldsymbol{d}_n]$ and the current state. Then we can define the following relation between the dimension of Ω_{t-1} and Ω_t by

$$|\Omega_t| \leq tR + (t-1)C + 2$$

where R is the number of regular states in Ω_{t-1}, and C is the number of critical states in Ω_{t-1}.

Note, the sequence of states spaces $\{\Omega_t\}_{t=1}^N$ is an extension for the sequence of states spaces $\{\Omega_t\}_{t=1}^N$ in Fu and Hsieh (2015). In the case of random permutations, for $t = 1, 2, \ldots, N$, we always insert an integer which is greater than all integers in $\pi(t-1)$. While, in the case of random permutations of arbitrary multi-sets at step t we insert an integer which is greater or equal to at least one integer in $\pi(t-1)$. The vector $[\boldsymbol{x}_{z'}(t-1)]$ provides the required information about the endpoints of the leftmost increasing subsequences in $\pi(t-1)$(the position and the value of these endpoints).

The above definitions and Proposition 4.1 provide the following proposition, which gives the transition probabilities from $\Omega_{t-1} \to \Omega_t$ induced by inserting the $c(t)$-th copy of the integer "$k(t)$" in the i-th gap.

Proposition 4.2. *The transition probabilities from $\Omega_{t-1} \to \Omega_t$ are given as follows:*

1. *For the absorbing states $\emptyset$ and α,*

$$\Pr(Y_t = \emptyset \mid Y_{t-1} = \emptyset) = \Pr(Y_t = \alpha \mid Y_{t-1} = \alpha) = 1.$$

2. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]) = y_{t-1}$ is a lower critical state of Ω_{t-1}, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap then

$$\Pr(Y_t = \emptyset \mid Y_{t-1} = y_{t-1}) =$$

$$\begin{cases} \frac{a_{b'}(t-1)}{t} & \text{if } c(t) = 1, \\ \frac{a_{b'}(t-1)}{t} & \text{if } c(t) > 1, \text{ and } x_{z'}(t-1) = a_{b'}(t-1), \\ \frac{x_{z'}(t-1)}{t} & \text{if } c(t) > 1, \text{ and } x_{z'}(t-1) > a_{b'}(t-1), \\ 1 & \text{if } c(t) > 1, \text{ and } x_{z'}(t-1) < a_{b'}(t-1), \\ 0 & \text{otherwise.} \end{cases}$$

For $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) = y_t$,

$$\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) =$$

$$\begin{cases} \frac{1}{t} & \text{if } c(t) = 1, b = b' + 1, \text{ and } z = b', \\ \frac{1}{t} & \text{if } c(t) > 1, x_{z'}(t-1) \geq a_{b'}(t-1), b = b' + 1, \text{ and } z = z', \\ 0 & \text{otherwise.} \end{cases}$$

3. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]) = y_{t-1}$ is an upper critical state of Ω_{t-1}, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap then

$$\Pr(Y_t = \alpha \mid Y_{t-1} = y_{t-1}) =$$

$$\begin{cases} \frac{t - a_{b'}(t-1)}{t} & \text{if } c(t) = 1, \\ \frac{t - a_{b'}(t-1)}{t} & \text{if } c(t) > 1, \text{ and } x_{z'}(t-1) = a_{b'}(t-1), \\ \frac{t - x_{z'}(t-1)}{t} & \text{if } c(t) > 1, \text{ and } x_{z'}(t-1) > a_{b'}(t-1), \\ 0 & \text{otherwise.} \end{cases}$$

For $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) = y_t$,

$$\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \begin{cases} \frac{1}{t} & \text{if } c(t) = 1, b = b', \text{ and } z = b', \\ \frac{1}{t} & \text{if } c(t) > 1, b = b', \text{ and } z = z', \\ 0 & \text{otherwise.} \end{cases}$$

4. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]) = y_{t-1}$ be a regular state in Ω_{t-1}, $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) = y_t$, and the 1st copy of the integer "$k(t)$" is inserted into the i-th gap then

$\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) =$

$$\begin{cases} \frac{1}{t} & \text{if } b = b', z = b', a_1(t) = 1, a_l(t) = a_l(t-1) + 1 \text{ for} \\ & l = 2, \ldots, b', \text{ and } x_r(t) = a_r(t-1) + 1 \text{ for } r = 1, \ldots, b', \\ \frac{1}{t} & \text{if } b = b', z = b', \text{ for } j = 2, \ldots, b', a_l(t) = x_l(t) = a_l(t-1) \\ & \text{for } l = 1, \ldots, j-1, a_j(t) = i, \text{ but } x_j(t) = a_j(t-1) + 1, \\ & \text{for } a_{j-1}(t-1) < i \le a_j(t-1), \text{ and } a_l(t) = x_l(t) = \\ & a_l(t-1) + 1, \text{ for } l = j+1, \ldots, b', \\ \frac{1}{t} & \text{if } b = b' + 1, z = b', a_l(t) = x_l(t) = a_l(t-1) \text{ for} \\ & l = 1, \ldots, b' \text{ and } a_{b'+1}(t) = i \text{ for } i > a_{b'}(t-1), \\ 0 & \text{otherwise.} \end{cases}$$

5. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)]) = y_{t-1}$ be a regular state in Ω_{t-1}, $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)]) = y_t$, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap, such that $c(t) > 1$, then:

   i) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, if $b = b', z = z', a_1(t) = 1, a_l(t) = a_l(t-1) + 1$ for $l = 2, 3, \ldots, b'$, and $x_r(t) = x_r(t-1) + 1$ for $r = 1, 2, \ldots, z'$.

ii) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, $b = b', z = z'$, *and for* $j = 2, 3, \ldots, b'$ *let*

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, 2, \ldots, j-1, \\ i & \text{for } l = j \text{ and } a_{j-1}(t-1) = \\ & x_{j-1}(t-1) \text{ or } x_{b'-1}(t-1) < i, \\ a_l(t-1)+1 & \text{for } l = j, a_{j-1}(t-1) \neq x_{j-1}(t-1) \\ & \text{and } x_{b'-1}(t-1) > i \text{ or } z' < b'-1, \\ a_l(t-1)+1 & \text{for } l = j+1, \ldots, b'. \end{cases}$$

and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } r \leq m, \\ x_r(t-1)+1 & \text{for } r > m; \end{cases}$$

where

$$m = \begin{cases} 0 & \text{for } i \leq x_1(t-1), \\ \max(d : 1 \leq d \leq z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \leq \\ & x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1). \end{cases}$$

iii) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{t - a_{b'}(t-1)}{t}$, *if* $b = b', z = z', x_{z'}(t-1) < a_{b'}(t-1)$, $a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b'$, *and* $x_r(t) = x_r(t-1)$ *for* $r = 1, 2, \ldots, z'$.

iv) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, *if* $b = b'+1, z = z', x_{z'}(t-1) = a_{b'}(t-1)$, $a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b'$, $a_{b'+1}(t) = i$ *for* $i > a_{b'}(t-1)$, *and* $x_r(t) = x_r(t-1)$ *for* $r = 1, 2, \ldots, z'$.

v) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, *if* $b = b'+1, z = z', x_{z'}(t-1) > a_{b'}(t-1)$, $a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b'$, $a_{b'+1}(t) = i$ *for*

57

$i > x_{z'}(t-1)$ *and*

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } r \leq m, \\ x_r(t-1)+1 & \text{for } r > m; \end{cases}$$

where

$$m = \begin{cases} 0 & \text{for } i \leq x_1(t-1), \\ \max(d : 1 \leq d \leq z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \leq \\ & x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1). \end{cases}$$

vi) $\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{x_{z'}(t-1) - a_{b'}(t-1)}{t}$, *if* $b = b', z =$

$z', x_{z'}(t-1) > a_{b'}(t-1), i \leq x_{z'}(t-1), a_l(t) = a_l(t-1)$ *for*

$l = 1, 2, \ldots, b'$, *and*

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } r \leq m, \\ x_r(t-1)+1 & \text{for } r > m; \end{cases}$$

where

$$m = \begin{cases} 0 & \text{for } i \leq x_1(t-1), \\ \max(d : 1 \leq d \leq z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \leq \\ & x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1). \end{cases}$$

Proof: If Y_{t-1} is a regular state of Ω_{t-1}, then the length of the longest increasing subsequence either increases by one or stays the same following the definition of both $[\boldsymbol{a}_b(t)]$ and $[\boldsymbol{x}_z(t)]$, and the relations specified in Proposition 4.1 with probability $\frac{1}{t}$, which proves part (4).

For the critical states, if Y_{t-1} is a lower critical state of Ω_{t-1}, then Y_{t-1} will always truncate to $\emptyset$ if inserting the $c(t)$-th copy of the integer "$k(t)$" did not create an increasing subsequence of length $b' + 1$. Otherwise, Y_{t-1} will never truncate, which proves part (2).

If Y_{t-1} is an upper critical state of Ω_{t-1}, then Y_{t-1} will always truncate to α if inserting the $c(t)$-th copy of the integer "$k(t)$" create an increasing subsequence of length v. Otherwise, Y_{t-1} will never truncate, which proves part (3). $\qquad\square$

Theorem 4.1. *Given N, $2 \le v \le N$, and ζ_1,*

$$\Pr(L_N([\boldsymbol{d}_n], \pi) < v) = \Pr(Y_N = \emptyset \mid \zeta_1) = \zeta_1 \left(\prod_{t=2}^{N} M_t \right) (1 \quad 0)^{\mathsf{T}},$$

and

$$\left| \{ \pi \in \Pi[\boldsymbol{d}_n] : L_N([\boldsymbol{d}_n], \pi) < v \} \right| = \frac{N!}{s_1! s_2! \dots s_n!} \zeta_1 \left(\prod_{t=2}^{N} M_t \right) (1 \quad 0)^{\mathsf{T}},$$

where $\left| \{ \pi \in \Pi[\boldsymbol{d}_n] : L_N([\boldsymbol{d}_n], \pi) < v \} \right|$ is the number of permutations where the length of the longest increasing subsequences is less than v.

Example 4.1. *Let $N = 5$, $n = 3$, and $v = 3$, with $D([\boldsymbol{d}_3]) = \{1, 2, 2, 3, 3\}$. Then the state spaces $\{\Omega_t\}_{t=1}^{5}$ associated with the $\{Y_t\}_{t=1}^{5}$ are as follows:*

If $t = 1$, then $\underline{B}_1 = 1$, $\bar{B}_1 = 1$, and $\Omega_1 = \{\emptyset, ([1], 0), \alpha\}$.

If $t = 2$, then $\underline{B}_2 = 1$, $\bar{B}_2 = 2$, and $\Omega_2 = \{\emptyset, ([1], [2]), ([1, 2], [1]), \alpha\}$.

If $t = 3$, then $\underline{B}_3 = 1$, $\bar{B}_3 = 2$, and

$$\Omega_3 = \{\emptyset, ([1], [3]), ([1, 2], [1]), ([1, 3], [2]), \alpha\}.$$

If $t = 4$, then $\underline{B}_4 = 2$, $\bar{B}_4 = 2$, and

$$\Omega_4 = \{\emptyset, ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 2], [1, 4]), ([1, 3], [1]),$$

$$([1, 3], [2, 3]), ([1, 3], [1, 4]), ([1, 4], [1]), ([1, 4], [2, 4]), \alpha\}.$$

If $t = 5$, then $\underline{B}_5 = 3$, $\bar{B}_5 = 2$, and $\Omega_5 = \{\emptyset, \alpha\}$. Then the transition matrices are given as follows:

$$M_2 = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad M_3 = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$M_4 = \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad M_5 = \frac{1}{5} \begin{bmatrix} 5 & 0 \\ 5 & 0 \\ 3 & 2 \\ 4 & 1 \\ 5 & 0 \\ 3 & 2 \\ 4 & 1 \\ 5 & 0 \\ 4 & 1 \\ 0 & 5 \end{bmatrix}.$$

Then for $\zeta_1 = (0, 1, 0)$

$$\Pr(L_5([\boldsymbol{d}_3], \pi) < 3) = \zeta_1 \left(\prod_{t=2}^{5} M_t \right) (1 \quad 0)^{\mathsf{T}} = 0.6\bar{3},$$

and

$$|\{\pi \in \Pi[\boldsymbol{d}_3] : L_5([\boldsymbol{d}_3], \pi) < 3\}| = \frac{5!}{1!2!2!} \zeta_1 \left(\prod_{t=2}^{5} M_t \right) (1 \quad 0)^{\mathsf{T}}$$

$$= 30 \times 0.6\bar{3}$$

$$= 19.$$

Example 4.2. *Let $N = 5$, $n = 5$, and $v = 4$, with $D([\boldsymbol{d}_5]) = \{1, 2, 3, 4, 5\}$. Then the state spaces $\{\Omega_t\}_{t=1}^{5}$ associated with the $\{Y_t\}_{t=1}^{5}$ are as follows:*

If $t = 1$, then $\underline{B}_1 = 1$, $\bar{B}_1 = 1$, and $\Omega_1 = \{\emptyset, ([1], 0), \alpha\}$.

If $t = 2$, then $\underline{B}_2 = 1$, $\bar{B}_2 = 2$, and $\Omega_2 = \{\emptyset, ([1], [2]), ([1, 2], [1]), \alpha\}$.

If $t = 3$, then $\underline{B}_3 = 2$, $\bar{B}_3 = 3$, and

$$\Omega_3 = \{\emptyset, ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 3], [1]), ([1, 3], [2, 3]),$$

$$([1, 2, 3], [1, 2]), \alpha\}.$$

If $t = 4$, then $\underline{B}_4 = 3$, $\bar{B}_4 = 3$, and

$$\Omega_4 = \{\emptyset, ([1, 2, 3], [1, 2]), ([1, 2, 3], [1, 2, 4]), ([1, 2, 4], [1, 2]),$$

$$([1, 2, 4], [1, 3, 4]), ([1, 3, 4], [1, 3]), ([1, 3, 4], [2, 3, 4]), \alpha\}.$$

If $t = 5$, then $\underline{B}_5 = 4$, $\bar{B}_5 = 3$, and $\Omega_5 = \{\emptyset, \alpha\}$. Then the transition matrices are given as follows:

$$M_2 = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad M_3 = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$
M_4 = \frac{1}{4}
\begin{bmatrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
2 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 4
\end{bmatrix},
\qquad
M_5 = \frac{1}{5}
\begin{bmatrix}
5 & 0 \\
3 & 2 \\
3 & 2 \\
4 & 1 \\
4 & 1 \\
4 & 1 \\
4 & 1 \\
0 & 5
\end{bmatrix}.
$$

Then for $\zeta_1 = (0, 1, 0)$

$$
\Pr(L_5([\boldsymbol{d}_5], \pi) < 4) = \zeta_1 \left(\prod_{t=2}^{5} M_t \right) \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathsf{T}}
$$

$$
= 0.858\bar{3},
$$

and

$$
|\{\pi \in \Pi[\boldsymbol{d}_5] : L_5([\boldsymbol{d}_5], \pi) < 4\}| = 5!\zeta_1 \left(\prod_{t=2}^{5} M_t \right) \begin{pmatrix} 1 & 0 \end{pmatrix}^{\mathsf{T}}
$$

$$
= 120 \times 0.858\bar{3}
$$

$$
= 103.
$$

Example 4.3. *Let $N = 5$, $n = 3$, and $v = 3$, with $D([\boldsymbol{d}_3]) = \{1, 2, 3, 3, 3\}$. Then the state spaces $\{\Omega_t\}_{t=1}^{5}$ associated with the $\{Y_t\}_{t=1}^{5}$ are as follows:*

If $t = 1$, then $\underline{B}_1 = 1$, $\bar{B}_1 = 1$, and $\Omega_1 = \{\emptyset, ([1], 0), \alpha\}$.

If $t = 2$, then $\underline{B}_2 = 1$, $\bar{B}_2 = 2$, and $\Omega_2 = \{\emptyset, ([1], [2]), ([1, 2], [1]), \alpha\}$.

If $t = 3$, then $\underline{B}_3 = 1$, $\bar{B}_3 = 2$, and

$$
\Omega_3 = \{\emptyset, ([1], [2]), ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 3], [1]),
$$

$$
([1, 3], [2, 3]), \alpha\}.
$$

If $t = 4$, then $\underline{B}_4 = 2$, $\bar{B}_4 = 2$, and

$$\Omega_4 = \{\emptyset, ([1,2],[1]), ([1,2],[1,4]), ([1,3],[1]), ([1,3],[2]),$$

$$([1,3],[2,4]), ([1,4],[2]), ([1,4],[3,4]), \alpha\}.$$

If $t = 5$, then $\underline{B}_5 = 3$, $\bar{B}_5 = 2$, and $\Omega_5 = \{\emptyset, \alpha\}$. Then the transition matrices are given as follows:

$$M_2 = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \qquad M_3 = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$M_4 = \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \qquad M_5 = \frac{1}{5} \begin{bmatrix} 5 & 0 \\ 5 & 0 \\ 4 & 1 \\ 5 & 0 \\ 5 & 0 \\ 4 & 1 \\ 5 & 0 \\ 4 & 1 \\ 0 & 5 \end{bmatrix}.$$

Then for $\zeta_1 = (0, 1, 0)$

$$\Pr(L_5([\boldsymbol{d}_3], \pi) < 3) = \zeta_1 \left(\prod_{t=2}^{5} M_t \right) (1 \quad 0)^\mathsf{T}$$

$$= 0.7,$$

and

$$|\{\pi \in \Pi[\boldsymbol{d}_3] : L_5([\boldsymbol{d}_3], \pi) < 3\}| = \frac{5!}{1!2!2!} \zeta_1 \left(\prod_{t=2}^{5} M_t \right) (1 \quad 0)^\mathsf{T}$$

$$= 20 \times 0.7$$

$$= 14.$$

63

Corollary 4.1. *Let $[\boldsymbol{d}_n] = [\boldsymbol{1}_n]$, then the vector $[\boldsymbol{x}_z(t)]$ depends only on the vector $[\boldsymbol{a}_{b'}(t-1)]$ and the insertion position i for $i = 1, 2, \ldots, t$.*

Proof: See Proposition 4.1 when $c(t) = 1$. $\qquad\qquad\qquad\qquad$ $\square$

For $c(t) = 1$, the vector $[\boldsymbol{x}_z(t)]$ does not depend on $[\boldsymbol{x}_{z'}(t-1)]$ and is a function of $[\boldsymbol{a}_{b'}(t-1)]$ for all $t = 1, 2, \ldots, N$, hence the transitions probabilities from $\Omega_{t-1} \to \Omega_t$ can be calculated based only on $[\boldsymbol{a}_{b'}(t-1)]$ and the insertion position i. Without loss of generality, we require only the vector $[\boldsymbol{a}_b(t)]$ to maintain the smallest state space Ω_t, in this case our results will be identical to Fu and Hsieh (2015) results.

Example 4.4. *Let $n = 6$ and $v = 3$, with $[\boldsymbol{1}_6]$. The state spaces associated with the induced Markov chain $\{Y_t\}_{t=1}^{6}$ where $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)])$ for $t = 1, 2, \ldots, 6$ are given by $\Omega_= \{\emptyset, ([1], 0), \alpha\}$, $\Omega_2 = \{\emptyset, ([1], [2]), ([1, 2], [1]), \alpha\}$,*

$$\Omega_3 = \{\emptyset, ([1], [2]), ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 3], [1]), ([1, 3], [2, 3]),$$

$$\alpha\},$$

$$\Omega_4 = \{\emptyset, ([1], [2]), ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 2], [1, 4]), ([1, 3], [1]),$$

$$([1, 3], [1, 4]), ([1, 3], [2, 3]), ([1, 4], [1]), ([1, 4], [2, 4]), \alpha\},$$

$$\Omega_5 = \{\emptyset, ([1, 2], [1]), ([1, 2], [1, 3]), ([1, 2], [1, 4]), ([1, 2], [1, 5]),$$

$$([1, 3], [1]), ([1, 3], [1, 4]), ([1, 3], [1, 5]), ([1, 3], [2, 3]),$$

$$([1, 4], [1]), ([1, 4], [1, 5]), ([1, 4], [2, 4]), ([1, 5], [1]),$$

$$([1, 5], [2, 5]), \alpha\},$$

and $\Omega_6 = \{\emptyset, \alpha\}$. With transition matrices

$$M_2 = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad M_3 = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$M_4 = \frac{1}{4}\begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix},$$

$$M_5 = \frac{1}{5}\begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix},$$

and

$$M_6 = \frac{1}{6} \begin{bmatrix} 6 & 0 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 2 & 4 \\ 3 & 3 \\ 3 & 3 \\ 3 & 3 \\ 3 & 3 \\ 4 & 2 \\ 4 & 2 \\ 4 & 2 \\ 5 & 1 \\ 5 & 1 \\ 0 & 6 \end{bmatrix}.$$

Since we only need $[\boldsymbol{a}_{b'}(t-1)]$ to obtain the transition matrices $\{M_t\}_{t=1}^6$, these state spaces can be reduced to the state spaces associated with Fu and Hsieh (2015) induced Markov chain $\{Y_t\}_{t=1}^6$ where $Y_t = [\boldsymbol{a}_b(t)]$ for $t = 1, 2, \ldots, 6$. In this case the state spaces are given by

$$\Omega_1 = \{\emptyset, [1], \alpha\}, \Omega_2 = \{\emptyset, [1], [1,2], \alpha\}, \Omega_3 = \{\emptyset, [1], [1,2], [1,3], \alpha\},$$

$$\Omega_4 = \{\emptyset, [1], [1,2], [1,3], [1,4], \alpha\}, \Omega_5 = \{\emptyset, [1,2], [1,3], [1,4], [1,5], \alpha\},$$

and $\Omega_6 = \{\emptyset, \alpha\}$. With transition matrices

$$M_2 = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \quad M_3 = \frac{1}{3} \begin{bmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

66

$$M_4 = \frac{1}{4} \begin{bmatrix} 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}, \quad M_5 = \frac{1}{5} \begin{bmatrix} 5 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 3 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix},$$

and

$$M_6 = \frac{1}{6} \begin{bmatrix} 6 & 0 \\ 2 & 4 \\ 3 & 3 \\ 4 & 2 \\ 5 & 1 \\ 0 & 6 \end{bmatrix}.$$

In both cases, for $\zeta_1 = (0, 1, 0)$

$$\Pr(L_6([\mathbf{1}_6], \pi) < 3) = \zeta_1 \left(\prod_{t=2}^{6} M_t \right) (1 \quad 0)^{\mathsf{T}}$$

$$= 0.18\bar{3},$$

and

$$|\{\pi \in \Pi[\mathbf{1}_6] : L_6([\mathbf{1}_6], \pi) < 3\}| = 6! \zeta_1 \left(\prod_{t=2}^{6} M_t \right) (1 \quad 0)^{\mathsf{T}}$$

$$= 720 \times 0.18\bar{3}$$

$$= 132.$$

And this agrees with Fu and Hsieh (2015).

4.2   The unconditional case

Let $L_N(\prod(n))$ denote the length of the longest increasing subsequence in a random sequence of length N generated from $\{1, 2, \ldots, n\}$ with

equal selection probabilities. The exact distribution of $L_N(\prod(n))$ can be calculated as follows:

$$\Pr(L_N(\prod(n)) < v) = \sum_{[d_n]} \Pr(L_N([d_n], \pi) < v \mid [d_n]) \times \Pr([d_n]) \quad (4.1)$$

which corresponds to averaging over all possible specifications $[d_n]$. For the purpose of this section, let the transitions probabilities for a given specification $[d_n]$ be denoted by

$$P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) \text{ for } t = 1, 2, \ldots, N.$$

Definition 4.5. *For a given permutation $\pi(t)$, and $[d_n]$, let $h_t \in \{1, \ldots, n\}$ be the largest integer among all integers in $\pi(t)$ (i.e. $h_t = k(t)$).*

Given $[d_n]$, let $\langle([a_{b'}(t-1)], [x_{z'}(t-1)], h_{t-1}), (t, i)\rangle$ denote the mapping induced by inserting the $c(t)$-th copy of the integer $k(t)$ into the i-th gap in $\pi(t-1)$. During the insertion procedure keeping track of h_t reduces the state space as well as the umber of terms in the summation in Equation (4.1). The following proposition specifies the relation between $([a_{b'}(t-1)], [x_{z'}(t-1)], h_{t-1})$ and $([a_b(t)], [x_z(t)], h_t)$.

Proposition 4.3. *For given $[d_n]$, and $([a_{b'}(t-1)], [x_{z'}(t-1)], h_{t-1})$, for $1 \leq b' \leq t-1$, $1 \leq z' \leq b'$, and $h_{t-1} \in \{1, 2, \ldots, n\}$.*

(1) If $i = 1$, and $c(t) = 1$, then

$$\langle([a_{b'}(t-1)], [x_{z'}(t-1)], h_{t-1}), (t, i)\rangle = ([a_{b'}(t)], [x_{b'}(t)], h_t),$$

where $h_t = k(t)$, $a_1(t) = 1$, $a_l(t) = a_l(t-1) + 1$ for $l = 2, \ldots, b'$, and $x_k(t) = a_k(t-1) + 1$ for $k = 1, \ldots, b'$.

68

(2) If $a_{j-1}(t-1) < i \le a_j(t-1)$ for some $j = 2, \ldots, b'$, and $c(t) = 1$, then

$$\langle ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t, i) \rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{b'}(t)], h_t),$$

where $h_t = k(t)$,

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, j-1, \\ i & \text{for } l = j, \\ a_l(t-1) + 1 & \text{for } l = j+1, \ldots, b'; \end{cases}$$

and

$$x_r(t) = \begin{cases} a_r(t-1) & \text{for } r = 1, \ldots, j-1, \\ a_r(t-1) + 1 & \text{for } r = j, \ldots, b'. \end{cases}$$

(3) If $a_{b'}(t-1) < i \le t$, and $c(t) = 1$, then

$$\langle ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t, i) \rangle = ([\boldsymbol{a}_{b'+1}(t)], [\boldsymbol{x}_{b'}(t)], h_t),$$

where $h_t = k(t)$,

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, b', \\ i & \text{for } l = b' + 1; \end{cases}$$

and

$$x_r(t) = a_r(t-1), \text{ for } r = 1, \ldots, b'.$$

(4) If $i = 1$, and $c(t) > 1$, then

$$\langle ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t, i) \rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)], h_t),$$

where $h_t = h_{t-1}$, $a_1(t) = 1$, $a_l(t) = a_l(t-1) + 1$ for $l = 2, \ldots, b'$, and $x_r(t) = x_r(t-1) + 1$ for $r = 1, \ldots, z'$.

69

(5) If $a_{j-1}(t-1) < i \le a_j(t-1)$ for $j = 2, \ldots, b'$, $c(t) > 1$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t,i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)], h_t),$$

where $h_t = h_{t-1}$,

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \ldots, j-1, \\ i & \text{for } l = j, \ \text{and } a_{j-1}(t-1) = x_{j-1}(t-1) \\ & \text{or } a_{j-1}(t-1) < x_{j-1}(t-1) < i, \\ a_l(t-1) + 1 & \text{for } l = j, a_{j-1}(t-1) < x_{j-1}(t-1), \\ & x_{j-1}(t-1) \ge i, \\ a_l(t-1) + 1 & \text{for } l = j+1, \ldots, b'; \end{cases}$$

and, for $r = 1, \ldots, z'$, let

$$m = \begin{cases} 0 & \text{for } i \le x_1(t-1), \\ \max(d : 1 \le d \le z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \\ & \le x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1); \end{cases}$$

and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } 1 \le r \le m, \\ x_r(t-1) + 1 & \text{for } m < r \le z'. \end{cases}$$

(6) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) > x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t,i)\rangle = ([\boldsymbol{a}_{b'}(t)], [\boldsymbol{x}_{z'}(t)], h_t),$$

where $h_t = h_{t-1}$,

$$a_l(t) = a_l(t-1), \ \text{for } l = 1, \ldots, b',$$

and

$$x_r(t) = x_r(t-1), \ \text{for } r = 1, \ldots, z'.$$

70

(7) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) = x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t,i)\rangle = ([\boldsymbol{a}_{b'+1}(t)], [\boldsymbol{x}_{z'}(t)], h_t),$$

where $h_t = h_{t-1}$,

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \dots, b', \\ i & \text{for } l = b' + 1; \end{cases}$$

and

$$x_r(t) = x_r(t-1), \ \text{for } r = 1, \dots, z'.$$

(8) If $a_{b'}(t-1) < i \le t$, $c(t) > 1$, and $a_{b'}(t-1) < x_{z'}(t-1)$, then

$$\langle([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}), (t,i)\rangle = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_{z'}(t)], h_t),$$

where $h_t = h_{t-1}$, $b = b'$ if $x_{z'}(t-1) \ge i$ or $b = b' + 1$ if $x_{z'}(t-1) < i$,

and

$$a_l(t) = \begin{cases} a_l(t-1) & \text{for } l = 1, \dots, b', \\ i & \text{for } l = b' + 1, \ \text{and } x_{z'}(t-1) < i; \end{cases}$$

and, for $r = 1, \dots, z'$, let

$$m = \begin{cases} 0 & \text{for } i \le x_1(t-1), \\ \max(d : 1 \le d \le z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \\ & \qquad \le x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1); \end{cases}$$

and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } 1 \le r \le m, \\ x_r(t-1) + 1 & \text{for } m < r \le z'. \end{cases}$$

71

Proof: Identical to Proposition 4.1. □

Definition 4.6. *Let $\underline{B}_t$, and $\bar{B}_t$ be defined as in definition 4.3. Then a state $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) \in \Omega_t$ is defined as a lower critical state if $b = v + t - N$ (i.e. $\underline{B}_t = max(1, v + t - N) = v + t - N$) and is defined as an upper critical state if $b = v - 1$ (i.e. $\bar{B}_t = min(k(t), v - 1) = v - 1$), otherwise it is referred as a regular state.*

Recall, that the critical states are the only states that could be absorbed. Now we are ready to defined the sequence of the state space $\{\Omega\}_{t=1}^N$ by

$$\Omega_t = \{([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) : b \geq \underline{B}_t, b \leq \bar{B}_t, 1 \leq z \leq b, \text{ and } h_t = 1, \ldots, n\}$$

$$\cup \{\emptyset, \alpha\},$$

where $\{\Omega_t\}_{t=1}^N$ is the sequence of the states spaces for a non-homogeneous Markov chain $\{Y_t\}_{t=1}^N$, where the state $\{\alpha\}$ represents an absorbing state such that if $Y_t \in \alpha$ then the length of the longest increasing subsequence is greater or equal to v, and the state $\{\emptyset\}$ represent an absorbing state such that if $Y_t \in \emptyset$ then the length of the longest increasing subsequence is less than v. We always start with $\Omega_1 = \{\emptyset, ([1], [0], 1), \ldots, ([1], [0], n), \alpha\}$, and since $\underline{B}_N > \bar{B}_N$, $\Omega_N = \{\emptyset, \alpha\}$

Definition 4.7. *For a given $([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t)$, let $H(t)$ denote the set of all specification $[\boldsymbol{s}_n] = [s_1, s_2, \ldots, s_n]$, such that $\sum_{i=1}^n s_i = N$, and $[\boldsymbol{s}_n]$ satisfies one of the following conditions:*

1. *If $[\boldsymbol{x}_z(t)] = 0$, then $[\boldsymbol{s}_n]$ contains zero copy of any integer less than integer "h_t", and at least one copy of integer "h_t".*

2. If $[\boldsymbol{x}_z(t)] \neq 0$, and $b' > z'$, then $[\boldsymbol{s}_n]$ contains at least one copy of $b' - 1$ distinct integers less than integer "h_t", and at least one copy of integer "h_t".

3. If $[\boldsymbol{x}_z(t)] \neq 0$, and $b' = z'$, then $[\boldsymbol{s}_n]$ contains at least one copy of b' distinct integers less than integer "h_t", and at least one copy of integer "h_t".

Proposition 4.4. *The transition probabilities for $\Omega_{t-1} \to \Omega_t$ for $t = 1, \ldots, N$, are given by*

$$\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \sum_{[\boldsymbol{s}_n] \in H(t-1)} P_{[\boldsymbol{s}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1})$$

$$\times \Pr([\boldsymbol{d}_n] = [\boldsymbol{s}_n] \mid [\boldsymbol{d}_n] \in H(t-1)),$$

$$(4.2)$$

which corresponds to summing over all specifications $[\boldsymbol{d}_n] \in H(t-1)$ in transitions from state Y_{t-1} to state Y_t, with

$$\Pr([\boldsymbol{d}_n] = [\boldsymbol{s}_n] \mid [\boldsymbol{d}_n] \in H(t-1)) = \frac{\binom{N}{s_1, s_2, \ldots, s_n}}{\sum_{[\boldsymbol{d}_n] \in H(t-1)} \binom{N}{d_1, d_2, \ldots, d_n}},$$

and the transition probabilities $P_{[\boldsymbol{d}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1})$ are given as follows:

1. *For the absorbing states $\emptyset$ and α,*

$$P_{[\boldsymbol{d}_n]}(Y_t = \emptyset \mid Y_{t-1} = \emptyset) = P_{[\boldsymbol{d}_n]}(Y_t = \alpha \mid Y_{t-1} = \alpha) = 1.$$

2. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}) = y_{t-1}$ is a lower critical state of Ω_{t-1}, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap then

$$P_{[\boldsymbol{d}_n]}(Y_t = \emptyset \mid Y_{t-1} = y_{t-1}) =$$

$$
\begin{cases}
\frac{a_{b'}(t-1)}{t} & \text{if } c(t) = 1, h_t > h_{t-1} \\
\frac{a_{b'}(t-1)}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, \text{ and } x_{z'}(t-1) = a_{b'}(t-1), \\
\frac{x_{z'}(t-1)}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, \text{ and } x_{z'}(t-1) > a_{b'}(t-1), \\
1 & \text{if } c(t) > 1, h_t = h_{t-1}, \\
& \text{and } x_{z'}(t-1) < a_{b'}(t-1), \\
0 & \text{otherwise.}
\end{cases}
$$

For $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) = y_t$,

$$P_{[\boldsymbol{d}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) =$$

$$
\begin{cases}
\frac{1}{t} & \text{if } c(t) = 1, h_t > h_{t-1}, b = b' + 1, \text{ and } z = b', \\
\frac{1}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, b = b' + 1, z = z', \text{ and} \\
& x_{z'}(t-1) \geq a_{b'}(t-1), \\
0 & \text{otherwise.}
\end{cases}
$$

3. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}) = y_{t-1}$ is an upper critical state of Ω_{t-1}, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap then

$$P_{[\boldsymbol{d}_n]}(Y_t = \alpha \mid Y_{t-1} = y_{t-1}) =$$

$$
\begin{cases}
\frac{t - a_{b'}(t-1)}{t} & \text{if } c(t) = 1, h_t > h_{t-1}, \\
\frac{t - a_{b'}(t-1)}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, \text{ and } x_{z'}(t-1) = a_{b'}(t-1), \\
\frac{t - x_{z'}(t-1)}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, \text{ and } x_{z'}(t-1) > a_{b'}(t-1), \\
0 & \text{otherwise.}
\end{cases}
$$

For $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) = y_t$,

$$P_{[\boldsymbol{d}_n]}(Y_t = y_1 \mid Y_{t-1} = y_{t-1}) =$$

$$\begin{cases} \frac{1}{t} & \text{if } c(t) = 1, h_t > h_{t-1}, b = b', \text{ and } z = b', \\ \frac{1}{t} & \text{if } c(t) > 1, h_t = h_{t-1}, b = b', \text{ and } z = z', \\ 0 & \text{otherwise.} \end{cases}$$

4. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}) = y_{t-1}$ be a regular state in Ω_{t-1}, and the 1st copy of the integer "$k(t)$" is inserted into the i-th gap. Then for $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) = y_t$ with $h_t > h_{t-1}$, we have

$$P_{[\boldsymbol{d}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) =$$

$$\begin{cases} \frac{1}{t} & \text{if } b = b', z = b', a_1(t) = 1, a_l(t) = a_l(t-1) + 1 \text{ for} \\ & l = 2, \ldots, b', \text{ and } x_r(t) = a_r(t-1) + 1 \text{ for } r = 1, \ldots, b', \\ \frac{1}{t} & \text{if } b = b', z = b', \text{ for } j = 2, \ldots, b', a_l(t) = x_l(t) = \\ & a_l(t-1) \text{ for } l = 1, \ldots, j-1, a_j(t) = i, \text{ but } x_j(t) = \\ & a_j(t-1) + 1, \text{ for } a_{j-1}(t-1) < i \le a_j(t-1), \\ & a_l(t) = x_l(t) = a_l(t-1) + 1, \text{ for } l = j+1, \ldots, b', \\ \frac{1}{t} & \text{if } b = b' + 1, z = b', a_l(t) = x_l(t) = a_l(t-1), \text{ for} \\ & l = 1, \ldots, b', \text{ and } a_{b'+1}(t) = i \text{ for } i > a_{b'}(t-1), \\ 0 & \text{otherwise.} \end{cases}$$

5. If $Y_{t-1} = ([\boldsymbol{a}_{b'}(t-1)], [\boldsymbol{x}_{z'}(t-1)], h_{t-1}) = y_{t-1}$ be a regular state in Ω_{t-1}, and the $c(t)$-th copy of the integer "$k(t)$" is inserted into the i-th gap, such that $c(t) > 1$. Then for $Y_t = ([\boldsymbol{a}_b(t)], [\boldsymbol{x}_z(t)], h_t) = y_t$ with $h_t = h_{t-1}$, we have:

   i) $P_{[\boldsymbol{d}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, if $b = b', z = z', h_t = h_{t-1}, a_1(t) = 1, a_l(t) = a_l(t-1) + 1$ for $l = 2, 3, \ldots, b'$, and $x_r(t) = x_r(t-1) + 1$ for $r = 1, 2, \ldots, z'$.

ii) $P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, $b = b', z = z'$, and for $j = 2, 3, \ldots, b'$ let

$$
a_l(t) = \begin{cases}
a_l(t-1) & \text{for } l = 1, 2, \ldots, j-1, \\
i & \text{for } l = j \text{ and } a_{j-1}(t-1) = \\
& x_{j-1}(t-1) \text{ or } x_{b'-1}(t-1) < i, \\
a_l(t-1)+1 & \text{for } l = j, a_{j-1}(t-1) \neq x_{j-1}(t-1) \\
& \text{and } x_{b'-1}(t-1) > i \text{ or } z' < b'-1, \\
a_l(t-1)+1 & \text{for } l = j+1, \ldots, b'.
\end{cases}
$$

and

$$
x_r(t) = \begin{cases}
x_r(t-1) & \text{for } r \leq m, \\
x_r(t-1)+1 & \text{for } r > m;
\end{cases}
$$

where

$$
m = \begin{cases}
0 & \text{for } i \leq x_1(t-1), \\
\max(d : 1 \leq d \leq z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \leq \\
& x_{z'}(t-1), \\
i & \text{for } i > x_{z'}(t-1).
\end{cases}
$$

iii) $P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{t - a_{b'}(t-1)}{t}$, *if* $b = b', z = z', x_{z'}(t-1) < a_{b'}(t-1), a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b'$, *and* $x_r(t) = x_r(t-1)$ *for* $r = 1, 2, \ldots, z'$.

iv) $P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, *if* $b = b'+1, z = z', x_{z'}(t-1) = a_{b'}(t-1), a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b', a_{b'+1}(t) = i$ *for* $i > a_{b'}(t-1)$, *and* $x_r(t) = x_r(t-1)$ *for* $r = 1, 2, \ldots, z'$.

v) $P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{1}{t}$, *if* $b = b'+1, z = z', x_{z'}(t-1) > a_{b'}(t-1), a_l(t) = a_l(t-1)$ *for* $l = 1, 2, \ldots, b', a_{b'+1}(t) = i$ *for*

76

$i > x_{z'}(t-1)$ and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } r \le m, \\ x_r(t-1)+1 & \text{for } r > m; \end{cases}$$

where

$$m = \begin{cases} 0 & \text{for } i \le x_1(t-1), \\ \max(d : 1 \le d \le z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \le \\ & \quad x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1). \end{cases}$$

vi) $P_{[d_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \frac{x_{z'}(t-1) - a_{b'}(t-1)}{t}$, if $b = b', z = z', x_{z'}(t-1) > a_{b'}(t-1), i \le x_{z'}(t-1), a_l(t) = a_l(t-1)$ for $l = 1, 2, \ldots, b'$, and

$$x_r(t) = \begin{cases} x_r(t-1) & \text{for } r \le m, \\ x_r(t-1)+1 & \text{for } r > m; \end{cases}$$

where

$$m = \begin{cases} 0 & \text{for } i \le x_1(t-1), \\ \max(d : 1 \le d \le z', x_d(t-1) < i) & \text{for } x_1(t-1) < i \le \\ & \quad x_{z'}(t-1), \\ i & \text{for } i > x_{z'}(t-1). \end{cases}$$

Proof: Identical to Proposition 4.2. $\qquad\square$

Theorem 4.2. *Given $N, n, 2 \le v \le N$, and the initial distribution*

$$\zeta_1' = \begin{bmatrix} 0 \\ \Pr([s_n] \mid s_1 \ge 1) \\ \Pr([s_n] \mid s_1 = 0, s_2 \ge 1) \\ \vdots \\ \Pr([s_n] \mid s_n = N) \end{bmatrix},$$

77

let

$$\Pr(L_N(\prod(n)) < v) = \Pr(Y_N = \emptyset \mid \zeta_1) = \zeta_1 \left(\prod_{t=2}^{N} M_t\right) (1 \quad 0)^{\mathsf{T}},$$

where $\{M_t\}_{t=1}^{N}$ is the transition matrices.

Example 4.5. *Let $N = 4$, $n = 3$, and $v = 3$. Then the state spaces $\{\Omega_t\}_{t=1}^{4}$ associated with the $\{Y_t\}_{t=1}^{4}$ are as follows:*

For $t = 1$, $\underline{B}_1 = 1$ and $\bar{B}_1 = 1$ and

$$\Omega_1 = \{\emptyset, ([1],[0],1), ([1],[0],2), ([1],[0],3), \alpha\}.$$

For $t = 2$, $\underline{B}_2 = 1$ and $\bar{B}_2 = 2$ and

$$\Omega_2 = \{\emptyset, ([1],[0],1), ([1],[0],2), ([1],[0],3), ([1],[2],2), ([1],[2],3),$$
$$([1,2],[1],2), ([1,2],[1],3), \alpha\}.$$

For $t = 3$, $\underline{B}_3 = 2$ and $\bar{B}_3 = 2$ and

$$\Omega_3 = \{\emptyset, ([1,2],[1],2), ([1,2],[1],3), ([1,3],[1],2), ([1,3],[1],3),$$
$$([1,3],[2],2), ([1,3],[2],3), ([1,2],[1,3],3), ([1,3],[2,3],3), \alpha\}.$$

For $t = 4$, $\underline{B}_4 = 3$ and $\bar{B}_4 = 2$ and $\Omega_4 = \{\emptyset, \alpha\}$. To illustrate the above definitions, the following table show how to obtain the transition matrix

M_2.

Y_1	$[\boldsymbol{d}_n] \in H(1)$	Y_2	$P_{[\boldsymbol{d}_n]}(Y_2 = Y_2 \mid Y_1 = y_1)$
$([1], [0], 1)$	$[4, 0, 0]$	$([1], [0], 1)$	$\frac{1}{65}$
	$[3, 1, 0]$	$([1], [0], 1)$	$\frac{4}{65}$
	$[3, 0, 1]$	$([1], [0], 1)$	$\frac{4}{65}$
	$[2, 2, 0]$	$([1], [0], 1)$	$\frac{6}{65}$
	$[2, 1, 1]$	$([1], [0], 1)$	$\frac{12}{65}$
	$[2, 0, 2]$	$([1], [0], 1)$	$\frac{6}{65}$
	$[1, 3, 0]$	$([1], [2], 2)$	$\frac{4}{65}$
	$[1, 2, 1]$	$([1], [2], 2)$	$\frac{12}{65}$
	$[1, 1, 2]$	$([1], [2], 2)$	$\frac{12}{65}$
	$[1, 0, 3]$	$([1], [2], 3)$	$\frac{4}{65}$
$([1], [0], 2)$	$[0, 4, 0]$	$([1], [0], 2)$	$\frac{1}{15}$
	$[0, 3, 1]$	$([1], [0], 2)$	$\frac{4}{15}$
	$[0, 2, 2]$	$([1], [0], 2)$	$\frac{6}{15}$
	$[0, 1, 3]$	$([1], [0], 3), ([1, 2], [1], 3)$	$\frac{2}{15}$
$([1], [0], 3)$	$[0, 0, 3]$	$([1], [0], 3)$	1

From the previous table, M_2 is given by

$$
M_2 =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{33}{65} & 0 & 0 & \frac{14}{65} & \frac{2}{65} & \frac{14}{65} & \frac{2}{65} & 0 \\
0 & 0 & \frac{11}{15} & 0 & 0 & \frac{2}{15} & 0 & \frac{2}{15} & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}.
$$

Similarly, we can obtain the following transition matrices:

$$M_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{17}{33} & \frac{6}{33} & \frac{2}{33} & \frac{6}{33} & \frac{2}{33} & 0 & 0 & 0 & 0 & 0 \\ \frac{7}{11} & 0 & \frac{2}{11} & 0 & \frac{2}{11} & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{11}{21} & 0 & \frac{3}{21} & 0 & \frac{3}{21} & \frac{4}{21} & 0 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{8}{21} & 0 & 0 & 0 & \frac{4}{21} & 0 & \frac{3}{21} & \frac{3}{21} & \frac{3}{21} \\ 0 & 0 & \frac{2}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$M_4 = \begin{bmatrix} 1 & 0 \\ \frac{11}{17} & \frac{6}{17} \\ 1 & 0 \\ \frac{15}{18} & \frac{3}{18} \\ 1 & 0 \\ \frac{13}{16} & \frac{3}{16} \\ 1 & 0 \\ \frac{3}{4} & \frac{1}{4} \\ \frac{3}{4} & \frac{1}{4} \\ 0 & 1 \end{bmatrix}.$$

If

$$\zeta_1 = (0, \frac{65}{81}, \frac{15}{81}, \frac{1}{81}, 0),$$

then

$$\Pr(L_4(\textstyle\prod(3)) < 3) = \zeta_1 \left(\prod_{t=2}^{4} M_t \right) (1 \quad 0)^{\mathsf{T}}$$

$$= 0.\bar{8}.$$

Example 4.6. *Let $N = 5$, $n = 3$, and $v = 3$. Then the state spaces $\{\Omega_t\}_{t=1}^5$ associated with the $\{Y_t\}_{t=1}^5$ are as follows:*

For $t = 1$, $\underline{B}_1 = 1$ and $\bar{B}_1 = 1$ and

$$\Omega_1 = \{\emptyset, ([1], [0], 1), ([1], [0], 2), ([1], [0], 3), \alpha\}.$$

For $t = 2$, $\underline{B}_2 = 1$ and $\bar{B}_2 = 2$ and

$$\Omega_2 = \{\emptyset, ([1], [0], 1), ([1], [0], 2), ([1], [0], 3), ([1], [2], 2), ([1], [2], 3),$$
$$([1, 2], [1], 2), ([1, 2], [1], 3), \alpha\}.$$

For $t = 3$, $\underline{B}_3 = 1$ and $\bar{B}_3 = 2$ and

$$\Omega_3 = \{\emptyset, ([1], [0], 1), ([1], [0], 2), ([1], [0], 3), ([1], [2], 2), ([1], [2], 3),$$
$$([1], [3], 2), ([1], [3], 3), ([1, 2], [1], 2), ([1, 2], [1], 3),$$
$$([1, 2], [1], 2), ([1, 2], [1], 3), ([1, 3], [1], 2), ([1, 3], [1], 3),$$
$$([1, 3], [2], 2), ([1, 3], [2], 3), ([1, 2], [1, 3], 3), ([1, 3], [2, 3], 3), \alpha\}.$$

For $t = 4$, $\underline{B}_4 = 2$ and $\bar{B}_4 = 2$ and

$$\Omega_4 = \{\emptyset, ([1, 2], [1], 2), ([1, 2], [1], 3), ([1, 3], [1], 2), ([1, 3], [1], 3),$$
$$([1, 4], [1], 2), ([1, 4], [1], 3), ([1, 3], [2], 2), ([1, 3], [2], 3),$$
$$([1, 4], [2], 2), ([1, 4], [2], 3), ([1, 4], [3], 2), ([1, 4], [3], 3),$$
$$([1, 2], [1, 4], 3), ([1, 3], [1, 4], 3), ([1, 2], [1, 3], 3), ([1, 3], [2, 3], 3),$$
$$([1, 3], [2, 4], 3), ([1, 4], [2, 4], 3), ([1, 4], [3, 4], 3), \alpha\}.$$

For $t = 5$, $\underline{B}_5 = 3$ and $\bar{B}_5 = 2$ and $\Omega_5 = \{\emptyset, \alpha\}$. We can obtain the

transition matrices following Proposition 4.5., for example

$$
M_2 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{262}{422} & 0 & 0 & \frac{75}{422} & \frac{5}{422} & \frac{75}{422} & \frac{5}{422} & 0 \\
0 & 0 & \frac{52}{62} & 0 & 0 & \frac{5}{62} & 0 & \frac{5}{62} & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \ and \ M_5 = \begin{bmatrix}
1 & 0 \\
\frac{53}{95} & \frac{42}{95} \\
1 & 0 \\
\frac{5}{7} & \frac{2}{7} \\
1 & 0 \\
\frac{13}{15} & \frac{2}{15} \\
1 & 0 \\
\frac{9}{13} & \frac{4}{13} \\
1 & 0 \\
\frac{17}{20} & \frac{3}{20} \\
1 & 0 \\
\frac{21}{25} & \frac{4}{25} \\
1 & 0 \\
\frac{4}{5} & \frac{1}{5} \\
\frac{4}{5} & \frac{1}{5} \\
\frac{3}{5} & \frac{2}{5} \\
\frac{3}{5} & \frac{2}{5} \\
\frac{4}{5} & \frac{1}{5} \\
\frac{4}{5} & \frac{1}{5} \\
\frac{4}{5} & \frac{1}{5} \\
0 & 1
\end{bmatrix}
$$

If

$$
\zeta_1 = (0, \frac{211}{243}, \frac{31}{243}, \frac{1}{243}, 0)
$$

then

$$\Pr(L_5(\prod(3)) < 3) = \zeta_1 \left(\prod_{t=2}^{5} M_t \right) (1 \quad 0)^\mathsf{T}$$

$$= 0.79.$$

Proposition 4.4 can be generalized for non-iid trials. For $\{1, 2, \ldots, n\}$ with unequal selection probabilities $\{p_1, p_2, \ldots, p_n\}$ the transitions probabilities are given in the following proposition.

Proposition 4.5. *The transition probabilities for $\Omega_{t-1} \to \Omega_t$ for $t = 1, \ldots, N$, are given by*

$$\Pr(Y_t = y_t \mid Y_{t-1} = y_{t-1}) = \sum_{[\boldsymbol{s}_n] \in H(t-1)} P_{[\boldsymbol{s}_n]}(Y_t = y_t \mid Y_{t-1} = y_{t-1})$$

$$\times \Pr([\boldsymbol{d}_n] = [\boldsymbol{s}_n] \mid [\boldsymbol{d}_n] \in H(t-1)), \tag{4.3}$$

which corresponds to summing over all specifications $[\boldsymbol{d}_n] \in H(t-1)$ in transitions from state Y_{t-1} to state Y_t, with

$$\Pr([\boldsymbol{d}_n] = [\boldsymbol{s}_n] \mid [\boldsymbol{d}_n] \in H(t-1)) = \frac{\binom{N}{s_1, s_2, \ldots, s_n} \prod_{i=1}^{n} p_i^{s_i}}{\sum_{[\boldsymbol{d}_n] \in H(t-1)} \binom{N}{d_1, d_2, \ldots, d_n} \prod_{i=1}^{n} p_i^{d_i}}.$$

Chapter 5

Summary and future work

In this dissertation, we studied the exact distribution for some run-related statistics using the finite Markov chain imbedding technique. In Chapter 1 we gave an introduction about the distribution theory of runs and patterns, and discussed some of the proposed methods for calculating the distribution of run-related statistics. One of these methods is the finite Markov chain imbedding technique which was proposed by Fu and Koutras (1994). We introduced some preliminary results regarding the finite Markov chain imbedding technique in Chapter 2.

In Chapter 3, we studied the exact distribution of the total number of adjacent and adjacent consecutive increasing subsequences in sequence using the finite Markov chain imbedding technique. To this end we defined a state space Ω for the induced Markov chain $\{Y_t\}_{t=1}^N$. Our approach give smaller state space than other proposed methods using the finite Markov chain imbedding technique. We can use results from Chapter 4 to calculate the exact distribution of the length of the longest increasing subsequence in sequences.

In Chapter 4, we focused on random permutations of arbitrary multi-sets. We proposed a new method to calculate the exact distribution of the length of the longest increasing subsequence in random permutation of arbitrary multi-sets using the finite Markov chain imbedding technique. First, we obtained the exact distribution of the length of the longest increasing subsequence in random permutation of arbitrary multi-sets given a specification $[\boldsymbol{d}_n]$. Then, we obtained the exact distribution of the length of the longest increasing subsequence in the set of all permutations of length N generated from $\{1, 2, \ldots, n\}$. Fu and Hsieh (2015) studied the exact distribution of the length of the longest increasing subsequence in random permutations but, to the best of our knowledge, little or no work has been done for the case of random permutation of arbitrary multi-sets.

We believe similar techniques may be used to find the exact distribution of the total number of adjacent and adjacent consecutive increasing subsequences in random permutation of arbitrary multi-sets. Another area of interest may be to study the properties of the distributions presented in this dissertation.

Bibliography

Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Courier Corporation. (Cited on page 9.)

Abramson, M. and Moser, W. (1967). Permutations without rising or falling ω-sequences. *The Annals of Mathematical Statistics* pages 1245–1254. (Cited on page 8.)

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes. *Annals of the Institute of Statistical Mathematics* **46,** 193–202. (Cited on page 2.)

Aldous, D. and Diaconis, P. (1995). Hammersley's interacting particle process and longest increasing subsequences. *Probability Theory and Related Fields* **103,** 199–213. (Cited on page 10.)

Aldous, D. and Diaconis, P. (1999). Longest increasing subsequences: from patience sorting to the Baik–Deift–Johansson theorem. *Bulletin of the American Mathematical Society* **36,** 413–432. (Cited on page 9.)

Baer, R. and Brock, P. (1968). Natural sorting over permutation spaces. *Mathematics of Computation* pages 385–410. (Cited on pages 9 and 10.)

Baik, J., Deift, P., and Johansson, K. (1999). On the distribution of the length of the longest increasing subsequence of random permutations. *Journal of the American Mathematical Society* **12,** 1119–1178. (Cited on pages 9 and 10.)

Boutsikas, M. V. and Koutras, M. V. (2000). A bound for the distribution of the sum of discrete associated or negatively associated random variables. *Annals of Applied Probability* pages 1137–1150. (Cited on page 3.)

Brook, D. and Evans, D. (1972). An approach to the probability distribution of CUSUM run length. *Biometrika* **59,** 539–549. (Cited on page 3.)

Carlitz, L. (1974). Permutations and sequences. *Advances in Mathematics* **14,** 92–120. (Cited on page 7.)

Carlitz, L. et al. (1964). Extended Bernoulli and Eulerian numbers. *Duke Mathematical Journal* **31,** 667–689. (Cited on pages 5 and 7.)

Chao, M. and Fu, J. C. (1989). A limit theorem of certain repairable systems. *Annals of the Institute of Statistical Mathematics* **41,** 809–818. (Cited on pages 3 and 4.)

Chao, M. and Fu, J. C. (1991). The reliability of a large series system under Markov structure. *Advances in Applied Probability* pages 894–908. (Cited on pages 3 and 4.)

Chao, M. and Lin, G. (1984). Economical design of large consecutive-k-out-of-n:F systems. *IEEE Transactions on Reliability* **33,** 411–413. (Cited on pages 3 and 4.)

Chen, L. H. (1975). An approximation theorem for convolutions of probability measures. *The Annals of Probability* **3,** 992–999. (Cited on page 3.)

Chryssaphinou, O., Papastavridis, S., and Tsapelas, T. (1993). On the number of overlapping success runs in a sequence of independent Bernoulli trials. In Bergum, G. E., Philippou, A. N., and Horadam, A. F., editors, *Applications of Fibonacci Numbers*, pages 103–112. Springer. (Cited on page 2.)

Crochemore, M. and Stefanov, V. T. (2003). Waiting time and complexity for matching patterns with automata. *Information Processing Letters* **87,** 119–125. (Cited on page 41.)

De Moivre, A. (1738). *The Doctrine of Chances.* H. Woodfall, London, 2 edition. (Cited on page 1.)

Deutsch, E., Hildebrand, A., and Wilf, H. S. (2003). Longest increasing subsequences in pattern-restricted permutations. *The Electronic Journal of Combinatorics* **9,** 1–8. (Cited on page 9.)

Dillon, J. and Roselle, D. (1969). Simon Newcomb's problem. *SIAM Journal on Applied Mathematics* **17,** 1086–1093. (Cited on pages 5 and 7.)

Dwass, M. (1973). The number of increases in a random permutation. *Journal of Combinatorial Theory, Series A* **15,** 192–199. (Cited on pages 5 and 8.)

Dymacek, W. M. and Roselle, D. (1978). Circular permutations by number of rises and successions. *Journal of Combinatorial Theory, Series A* **25,** 196–201. (Cited on page 8.)

Erdős, P. and Szekeres, G. (1935). A combinatorial problem in geometry. *Compositio Mathematica* **2,** 463–470. (Cited on page 9.)

Feller, W. (1968). *An introduction to probability theory and its applications*, volume 1. Wiley, New York, 3 edition. (Cited on page 15.)

Fredman, M. L. (1975). On computing the length of longest increasing subsequences. *Discrete Mathematics* **11,** 29–35. (Cited on page 38.)

Fu, J. C. (1986). Reliability of consecutive-k-out-of-n:F systems with (k-1)-step Markov dependence. *IEEE Transactions on Reliability* **35,** 602–606. (Cited on page 4.)

Fu, J. C. (1995). Exact and limiting distributions of the number of successions in a random permutation. *Annals of the Institute of Statistical Mathematics* **47,** 435–446. (Cited on pages 3, 6 and 8.)

Fu, J. C. and Hsieh, Y.-F. (2015). On the distribution of the length of the longest increasing subsequence in a random permutation. *Methodology*

and Computing in Applied Probability **17,** 489–496. (Cited on pages 3, 10, 44, 54, 64, 66, 67 and 86.)

Fu, J. C. and Hu, B. (1987). On reliability of a large consecutive-k-out-of-n:F system with (k-1)-step Markov dependence. *IEEE Transactions on Reliability* **36,** 75–77. (Cited on pages 3 and 4.)

Fu, J. C. and Koutras, M. (1994). Distribution theory of runs: a Markov chain approach. *Journal of the American Statistical Association* **89,** 1050–1058. (Cited on pages 3, 4, 16, 19, 20 and 85.)

Fu, J. C. and Lou, W. W. (2003). *Distribution theory of runs and patterns and its applications: a finite Markov chain imbedding approach.* World Scientific. (Cited on pages 39 and 40.)

Fu, J. C., Lou, W. W., and Wang, Y.-J. (1999). On the exact distributions of Eulerian and Simon Newcomb numbers associated with random permutations. *Statistics & Probability Letters* **42,** 115–125. (Cited on pages 3 and 7.)

Gibbons, J. D. (1985). *Nonparametric statistical inference.* Marcel Dekker, Inc., New York. (Cited on page 1.)

Giladi, E. and Keller, J. B. (1994). Eulerian number asymptotics. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* **445,** 291–303. (Cited on page 7.)

Godbole, A. P. (1990). Specific formulae for some success run distributions. *Statistics & Probability Letters* **10,** 119–124. (Cited on page 2.)

Gonnet, G. H. and Baeza-Yates, R. (1991). *Handbook of algorithms and data structures: in Pascal and C*, volume 2. Addison-Wesley Reading, MA. (Cited on page 38.)

Harris, B. and Park, C. (1994). A generalization of the Eulerian numbers with a probabilistic application. *Statistics & Probability Letters* **20,** 37–47. (Cited on page 7.)

Hirano, K. (1986). Some properties of the distributions of order k. In Philippou, A. N., Bergum, G. E., and Horadam, A. F., editors, *Fibonacci Numbers and Their Applications*, volume 42, pages 43–54. Reidel Dordrecht. (Cited on page 2.)

Hopcroft, J. E., Motwani, R., and Ullman, J. D. (2001). *Introduction to Automata theory, languages, and computation*, volume 24. Addison-Wesley, 2 edition. (Cited on page 41.)

Jackson, D. and Aleliunas, R. (1977). Decomposition based generating functions for sequences. *Canadian Journal of Mathematics* **29,** 971–1009. (Cited on page 8.)

Jackson, D. and Goulden, I. (1980). A formal calculus for the enumerative system of sequences-III. further developments. *Studies in Applied Mathematics* **62,** 113–143. (Cited on page 8.)

Jackson, D. and Goulden, I. (1981). Algebraic methods for permutations with prescribed patterns. *Advances in Mathematics* **42,** 113–135. (Cited on page 8.)

Jackson, D. and Reilly, J. (1976). Permutations with a prescribed number of p-runs. *Ars Combinatoria* **1,** 297–305. (Cited on page 8.)

Johansson, K. (1998). The longest increasing subsequence in a random permutation and a unitary random matrix model. *Mathematical Research Letters* **5,** 63–82. (Cited on page 10.)

Johnson, B. C. (2001). Distribution of increasing ℓ-sequences in a random permutation. *Methodology and Computing in Applied Probability* **3,** 35–49. (Cited on pages 5 and 8.)

Johnson, B. C. (2002). The distribution of increasing 2-sequences in random permutations of arbitrary multi-sets. *Statistics & Probability Letters* **59,** 67–74. (Cited on page 9.)

Johnson, B. C. and Fu, J. C. (2000). The distribution of increasing ℓ-sequences in random permutations: a Markov chain approach. *Statistics & Probability Letters* **49,** 337–344. (Cited on pages 3 and 8.)

Kaplansky, I. (1944). Symbolic solution of certain problems in permutations. *Bulletin of the American Mathematical Society* **50,** 906–914. (Cited on page 8.)

Kaplansky, I. (1945). The asymptotic distribution of runs of consecutive elements. *The Annals of Mathematical Statistics* **16,** 200–203. (Cited on page 8.)

Koutras, M. and Alexandrou, V. (1995). Runs, scans and urn model distributions: a unified Markov chain approach. *Annals of the Institute*

of Statistical Mathematics **47,** 743–766. (Cited on pages 18, 19, 20 and 37.)

Ling, K. (1988). On binomial distributions of order k. *Statistics & Probability Letters* **6,** 247–250. (Cited on page 2.)

Logan, B. F. and Shepp, L. A. (1977). A variational problem for random Young tableaux. *Advances in Mathematics* **26,** 206–222. (Cited on page 10.)

McDiarmid, C. (1998). Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, volume 16 of *Algorithms and Combinatorics*, pages 195–248. Springer Berlin Heidelberg. (Cited on page 38.)

Mood, A. M. (1940). The distribution theory of runs. *The Annals of Mathematical Statistics* **11,** 367–392. (Cited on page 1.)

Nicodeme, P., Salvy, B., and Flajolet, P. (2002). Motif statistics. *Theoretical Computer Science* **287,** 593–617. (Cited on page 40.)

Nicolas, J.-L. (1991). An integral representation for Eulerian numbers. In *Sets, Graphs and Numbers*, volume 60, pages 513–527. Colloquia Mathematica Societatis János Bolyai. (Cited on page 7.)

Odlyzko, A. and Rains, E. (2000). On longest increasing subsequences in random permutations. *Contemporary Mathematics* **251,** 439–452. (Cited on page 9.)

Philippou, A. N. and Makri, F. S. (1986). Successes, runs and longest runs. *Statistics & Probability Letters* **4,** 101–105. (Cited on page 2.)

Pilpel, S. (1990). Descending subsequences of random permutations. *Journal of Combinatorial Theory, Series A* **53,** 96–116. (Cited on page 10.)

Reilly, J. and Tanny, S. (1979). Counting successions in permutations. *Studies in Applied Mathematics* **61,** 73–81. (Cited on pages 5 and 8.)

Riordan, J. (1945). Permutations without 3-sequences. *Bulletin of the American Mathematical Society* **51,** 745–748. (Cited on page 8.)

Romik, D. (2014). The surprising mathematics of longest increasing subsequences. availables at: https://www.math.ucdavis.edu/ romik/download-book.php. (Cited on page 9.)

Roselle, D. et al. (1974). Graphs, quasisymmetry and permutations with restricted positions. *Duke Mathematical Journal* **41,** 41–50. (Cited on page 8.)

Roselle, D. P. (1968). Permutations by number of rises and successions. *Proceedings of the American Mathematical Society* pages 8–16. (Cited on pages 5, 6 and 8.)

Schensted, C. (1961). Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics* **13,** 179–191. (Cited on page 9.)

Stefanov, V. T., Robin, S., and Schbath, S. (2007). Waiting times for clumps of patterns and for structured motifs in random sequences. *Discrete Applied Mathematics* **155,** 868–880. (Cited on page 41.)

Stevens, W. (1939). Distribution of groups in a sequence of alternatives. *Annals of Eugenics* **9,** 10–17. (Cited on page 1.)

Sun, X. and Woodruff, D. P. (2007). The communication and streaming complexity of computing the longest common and increasing subsequences. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 336–345. Society for Industrial and Applied Mathematics. (Cited on page 38.)

Tanny, S. et al. (1973). A probabilistic interpretation of Eulerian numbers. *Duke Mathematical Journal* **40,** 717–722. (Cited on pages 5 and 7.)

Tanny, S. M. (1976). Permutations and successions. *Journal of Combinatorial Theory, Series A* **21,** 196–202. (Cited on pages 5 and 8.)

Ulam, S. M. (1961). Monte Carlo calculations in problems of mathematical physics. In Beckenbach, E. F., editor, *Modern Mathematics for the Engineers*, Second series, pages 261–281. McGraw Hill. (Cited on page 9.)

Vershik, A. M. and Kerov, S. V. (1977). Asymptotics of Plancherel measure of symmetrical group and limit form of Young tables. *Doklady Akademii Nauk SSSR* **233,** 1024–1027. (Cited on page 10.)

Wald, A. and Wolfowitz, J. (1940). On a test whether two samples are

from the same population. *The Annals of Mathematical Statistics* **11,** 147–162. (Cited on page 2.)

Walsh, J. E. (1962). *Handbook of Nonparametric Statistics.* Van Nostrand Reinhold. (Cited on page 1.)

Whitworth, W. A. (1901). *Choice and Chance: With 1000 Exercises.* Deighton Bell and Company; London George Bell and Sons. (Cited on page 8.)

Wolfowitz, J. (1942). Additive partition functions and a class of statistical hypotheses. *The Annals of Mathematical Statistics* **13,** 247–279. (Cited on page 8.)

Wolfowitz, J. (1944). Asymptotic distribution of runs up and down. *The Annals of Mathematical Statistics* **15,** 163–172. (Cited on page 8.)

Wu, T.-L. (2013). On finite Markov chain imbedding and its applications. *Methodology and Computing in Applied Probability* **15,** 453–465. (Cited on page 21.)