Ann. Inst. Statist. Math.
Vol. 52, No. 3, 415–425 (2000)

JOINT DISTRIBUTION OF RISES AND FALLS*

JAMES C. FU[1] AND W. Y. WENDY LOU[2]

[1] *Department of Statistics, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2*
[2] *Department of Biomathematical Sciences, Mount Sinai School of Medicine, New York, NY 10029-6574, U.S.A.*

(Received September 8, 1997; revised December 7, 1998)

Abstract. The marginal distributions of the number of rises and the number of falls have been used successfully in various areas of statistics, especially in non-parametric statistical inference. Carlitz (1972, *Duke Math. J.* **39**, 268–269) showed that the generating function of the joint distribution for the numbers of rises and falls satisfies certain complex combinatorial equations, and pointed out that he had been unable to derive the explicit formula for the joint distribution from these equations. After more than two decades, this latter problem remains unsolved. In this article, the joint distribution is obtained via the probabilistic method of finite Markov chain imbedding for random permutations. A numerical example is provided to illustrate the theoretical results and the corresponding computational procedures.

Key words and phrases: Eulerian and Simon Newcomb numbers, finite Markov chain imbedding, transition probability matrix.

1. Introduction

Let $S_N = \{1, \dots, 1, 2, \dots, 2, \dots, n, \dots, n\}$ be a set of N integers with specification $[s] = [s_1, \dots, s_n]$, where s_i (≥ 1) is the number of times the integer "i" occurs in S_N and $s_1 + s_2 + \cdots + s_n = N$. Let $\mathcal{H}(S_N) = \{\pi(N) = (\pi_1(N), \pi_2(N), \dots, \pi_N(N)) : \pi_i(N) \in S_N\}$ be the collection of all possible permutations having specification $[s]$. For $i = 1, \dots, N-1$, define the rises (increases) R_{iN}, falls (decreases) D_{iN}, and levels L_{iN} for the gaps among $\pi_1(N)$ to $\pi_N(N)$ by index functions:

$$R_{iN} = \begin{cases} 1 & \text{if } \pi_{i+1}(N) > \pi_i(N) \\ 0 & \text{otherwise,} \end{cases}$$

$$D_{iN} = \begin{cases} 1 & \text{if } \pi_{i+1}(N) < \pi_i(N) \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and}$$

$$L_{iN} = \begin{cases} 1 & \text{if } \pi_{i+1}(N) = \pi_i(N) \\ 0 & \text{otherwise,} \end{cases}$$

and by convention let the front end be a rise and the rear end be a fall ($R_{0N} \equiv 1$, and $D_{NN} \equiv 1$). Further, we define three random variables $R_N = \sum_{i=0}^{N-1} R_{iN}$, $D_N = \sum_{i=1}^{N} D_{iN}$, and $L_N = \sum_{i=1}^{N-1} L_{iN}$ as the numbers of rises, falls, and levels in a random

* This work was supported in part by the National Science and Engineering Research Council of Canada Grant NSERC A-9216, and by the NATO Collaborative Research Grant CRG-970278.

permutation $\pi(N) \in \mathcal{H}(S_N)$, respectively, and for every $\pi(N) \in \mathcal{H}(S_N)$, they satisfy the identity

$$R_N + D_N + L_N = N + 1. \tag{1.1}$$

Let $A([s], k)$, $k = 1, \ldots, n$, be the number of permutations in $\mathcal{H}(S_N)$ having exactly k rises. The numbers $A([s], k)$ are known as the Simon Newcomb numbers (Dillon and Roselle (1969)). If the specification $[s]$ is $[1, \ldots, 1]$, then the $A([s], k)$ are known as the Eulerian numbers (Carlitz (1964)). The distribution of the number of rises associated with random permutations $\pi(N) \in \mathcal{H}(S_N)$ can be directly defined through the Simon Newcomb numbers in the following way:

$$P(R_N = k) = A([s], k) \prod_{i=1}^{n} (s_i!)/N!. \tag{1.2}$$

Eulerian and Simon Newcomb numbers are probably the two most celebrated numbers associated with random permutations, and have been studied extensively in the literature on combinatorial analysis. The history of their development and application can be found in books by, for instance, MacMahon (1915), Riordon (1958), and David and Barton (1962). Today there remains a considerable amount of interest in generalizations and extensions of Eulerian and Simon Newcomb numbers (for example, Carlitz (1964, 1972, 1974), Tanny (1973), Takacs (1979), Nicolas (1992), Harris and Park (1994), Giladi and Keller (1994), and Fu *et al.* (1998)).

Carlitz (1972) studied the generating function of the joint distribution for the numbers of rises and falls. He showed that the generating function satisfies a set of complex implicit combinatorial equations (see Carlitz (1972), pp. 268–269). In that article, he also pointed out that an explicit formula for the joint distribution could not be obtained from his set of combinatorial equations. Further, Carlitz did not provide a numerical method to evaluate the joint probabilities. After more than two decades, the problem of finding the joint distribution remains unsolved. In this article, away from the traditional combinatorial approach, the joint distribution for the numbers of rises and falls is obtained using the probabilistic method of finite Markov chain imbedding for random permutations, as recently developed by Fu (1995) and Fu *et al.* (1998) to study the marginal distributions of successions, rises and falls. The key idea of this approach is to view the random vector of interest, (R_N, D_N), as a function (projection) of a simple finite Markov chain, and then its distribution can be expressed in terms of the transition probability matrices of the imbedded finite Markov chain.

This manuscript is organized in the following manner. In Section 2, the insertion procedure, the lemmas, and the finite Markov chain imbedding technique, are introduced. In Section 3, we study the joint distribution of the numbers of rises and falls. A numerical example and discussion are given in the concluding section.

2. Notation and preliminary results

Let S_N be a collection of N integers having specification $[s] = [s_1, \ldots, s_n]$, as defined in Section 1. For given t, let k_t be the integer such that

$$\sum_{i=1}^{k_t - 1} s_i < t \le \sum_{i=1}^{k_t} s_i \tag{2.1}$$

and

$$s_{k_t}(t) = t - \sum_{i=1}^{k_t-1} s_i. \tag{2.2}$$

For $t = 1, 2, \ldots, N$, we define a sequence of collections generated by S_N: $S_1 = \{1\}$, $S_2 = \{1, 1\}, \ldots, S_{s_1} = \{1, \ldots, 1\}, \ldots, S_t = S_{t-1} \cup \{k_t\}, \ldots$, and $S_N = \{1, \ldots, 1, \ldots, n, \ldots, n\}$. In words, the collection S_t contains the first t integers of S_N and has specification $[s]_t = [s_1, \ldots, s_{k_t}(t)]$. Let, for $t = 1, 2, \ldots, N$,

$$\begin{aligned} \mathcal{H}(S_t) &= \text{all the permutations generated by integers in } S_t \\ &= \{\pi(t) = (\pi_1(t), \ldots, \pi_t(t)) : \pi_i(t) \in S_t, i = 1, \ldots, t\}. \end{aligned} \tag{2.3}$$

From (2.1) and (2.2), for every t, either $k_t = k_{t-1}$ or $k_t = k_{t-1} + 1$, and $\{k_t\}_{t=1}^N$ is a non-decreasing sequence of integers.

In order to apply the finite Markov chain imbedding technique for random permutations, we adopt the insertion procedure introduced by Fu (1995) and Fu *et al.* (1998). The insertion procedure can be described as inserting integers $\{k_t\}$ one by one randomly into the gaps between integers (including the two end gaps), starting with the first s_1 "1"s, followed by s_2 "2"s, and continuing the insertion procedure until all the integers $\{k_t\}$ have been inserted. The permutation $\pi(t) \in \mathcal{H}(S_t)$ is a result of inserting the integer k_t into one of the t gaps of permutation $\pi(t-1) \in \mathcal{H}(S_{t-1})$. The sequence of triplets $\{(S_t, [s]_t, \mathcal{H}(S_t)) : t = 1, \ldots, N\}$ induced by S_N through the insertion procedure plays an important role in the application of the finite Markov chain imbedding technique.

Let us consider a random permutation $\pi = (1322331)$. There are eight gaps in the permutation (including the two end gaps). Given an integer $k = 3$, the eight gaps of the random permutation π can be classified into four types according to the given integer k as follows:

(I) gaps of rises, falls, and levels involving the integer k, i.e. gaps 2, 3, 5, 6, and 7,
(II) gaps of rises not involving the integer k, i.e. gap 1,
(III) gaps of falls not involving the integer k, i.e. gap 8,
(IV) gaps of levels not involving the integer k, i.e. gap 4.

Given integers k and t, and a random permutation π, we define six random variables: $R_t(k, \pi)$, $D_t(k, \pi)$, and $L_t(k, \pi)$ as the numbers of gaps of rises, falls, and levels involving the integer k in π respectively, and $R_t(\bar{k}, \pi)$, $D_t(\bar{k}, \pi)$, and $L_t(\bar{k}, \pi)$ as the numbers of gaps of rises, falls, and levels not involving the integer k in π, respectively.

Given S_N having specification $[s]$, the sequences $\{k_t\}$, $\{s_{k_t}(t)\}$, $\{S_t\}$, $\{[s]_t\}$, and $\{\mathcal{H}(S_t)\}$ are well defined, and the following lemmas hold.

LEMMA 2.1. *For every $1 \le t \le N$, the number of rises R_t, the number of falls D_t, and the number of levels L_t in a random permutation $\pi(t) \in \mathcal{H}(S_t)$ satisfy the following equations:*

(i) $R_t + D_t + L_t = t + 1$,
(ii) $1 \le R_t \le t - \max(s_1, s_2, \ldots, s_{k_t}(t)) + 1$,
(iii) $1 \le D_t \le t - \max(s_1, s_2, \ldots, s_{k_t}(t)) + 1$,
(iv) $0 \le L_t \le t - k_t$.

PROOF. For every $\pi(t) \in \mathcal{H}(S_t)$, there are $t+1$ gaps. Result (i) follows immediately from the fact that a gap is either a rise, a fall, or a level. Results (ii) and (iii) are due to Dillon and Roselle (1969). Result (iv) follows from

$$0 \le L_t \le (s_1 - 1) + (s_2 - 1) + \cdots + (s_{k_t}(t) - 1) = t - k_t. \qquad \square$$

LEMMA 2.2. *For every $\pi(t) \in \mathcal{H}(S_t)$, it follows that*
(i) $R_t(k_t, \pi(t)) + L_t(k_t, \pi(t)) = s_{k_t}(t)$,
(ii) $D_t(k_t, \pi(t)) + L_t(k_t, \pi(t)) = s_{k_t}(t)$,
(iii) $R_t(k_t, \pi(t)) = D_t(k_t, \pi(t))$.

PROOF. If $s_{k_t}(t) = 1$, there is no level associated with the integer k_t, and then $L_t(k_t, \pi(t)) = 0$. Since the integer k_t is the largest integer among all the integers in $\pi(t)$, the gap immediately preceding the integer k_t has to be a rise. It follows that $R(k_t, \pi(t)) = 1$ and result (i) is obviously true. For $s_{k_t}(t) \geq 2$, again since k_t is the largest integer in $\pi(t)$, the gap immediately preceding the integer k_t is either a rise or a level. Hence result (i) holds for all $s_{k_t} \geq 1$. By the same token, for all k_t, the gap immediately following the integer k_t has to be either a level or a fall. This completes the proof for result (ii). Result (iii) is a direct consequence of results (i) and (ii). □

For $t = 2, \ldots, N$, we define the index functions

$$I(t) = \begin{cases} 1 & \text{if} \quad k_t = k_{t-1}, \\ 0 & \text{if} \quad k_t = k_{t-1} + 1. \end{cases} \tag{2.4}$$

In the following, we show the relationships of the numbers of gaps associated with k_t and the numbers of gaps associated with k_{t-1}.

LEMMA 2.3. *For $t = 2, 3, \ldots, N$, and $\pi(t-1) \in \mathcal{H}(S_{t-1})$,*
(i) $R_{t-1} = R_{t-1}(k_{t-1}, \pi(t-1)) + R_{t-1}(\bar{k}_{t-1}, \pi(t-1))$,
$D_{t-1} = D_{t-1}(k_{t-1}, \pi(t-1)) + D_{t-1}(\bar{k}_{t-1}, \pi(t-1))$,
$L_{t-1} = L_{t-1}(k_{t-1}, \pi(t-1)) + L_{t-1}(\bar{k}_{t-1}, \pi(t-1))$,
(ii) $R_{t-1}(k_t, \pi(t-1)) + D_{t-1}(k_t, \pi(t-1)) + L_{t-1}(k_t, \pi(t-1))$
$= [2s_{k_{t-1}}(t-1) - L_{t-1}(k_{t-1}, \pi(t-1))]I(t)$,
(iii) $R_{t-1}(\bar{k}_t, \pi(t-1)) = I(t)R_{t-1}(\bar{k}_{t-1}, \pi(t-1)) + (1 - I(t))R_{t-1}$,
$D_{t-1}(\bar{k}_t, \pi(t-1)) = I(t)D_{t-1}(\bar{k}_{t-1}, \pi(t-1)) + (1 - I(t))D_{t-1}$,
$L_{t-1}(\bar{k}_t, \pi(t-1)) = I(t)L_{t-1}(\bar{k}_{t-1}, \pi(t-1)) + (1 - I(t))L_{t-1}$.

PROOF. Result (i) follows directly from the definitions. Result (ii) is a direct consequence of Lemma 2.2. If $I(t) = 0$ (i.e. $k_t = k_{t-1} + 1$), then all the rises in $\pi(t-1)$ do not involve the integer k_t (random permutation $\pi(t-1)$ contains no integer k_t), and $R_{t-1}(\bar{k}_t, \pi(t-1)) = R_{t-1}$. If $I(t) = 1$ (i.e. $k_t = k_{t-1}$), then the number of gaps of rises in $\pi(t-1)$ involving k_t is the same as that involving k_{t-1}, and $R_{t-1}(\bar{k}_t, \pi(t-1)) = R_{t-1}(\bar{k}_{t-1}, \pi(t-1))$. Hence the first part of (iii) holds. By the same token, the second and third parts of result (iii) are also true. □

Let X_N be a random variable or a random vector defined on $\mathcal{H}(S_N)$. To apply the finite Markov chain imbedding technique, we first give the basic definition:

DEFINITION 1. The random variable (or random vector) X_N defined on $\mathcal{H}(S_N)$ is *finite Markov chain imbeddable* if

(i) there exists a finite Markov chain $\{Y_t\}$ defined on a sequence of finite state spaces $\{\Omega_t\}$ having transition probability matrices $\{M_t\}$, $t = 1, 2, \ldots, N$, and initial probability ξ_0, and

(ii) there exists a partition $\{C_x : x = 0, 1, \ldots, l\}$ on the state space Ω_N such that for every $x = 0, 1, \ldots, l$,

$$P(X_N = x) = P(Y_N \in C_x \mid \xi_0).$$

If the random variable X_N is finite Markov chain imbeddable, then it follows from the Chapman-Kolmogorov equation that the exact distribution, moments, and generating functions can be obtained, respectively, from the following equations (see Fu and Koutras (1994) for details):

$$P(X_N = x) = \xi_0 \left(\prod_{t=1}^{N} M_t \right) U'(C_x), \qquad x = 0, 1, \ldots, l, \tag{2.5}$$

where $U(C_x) = \sum_{a \in C_x} U(a)$ and $U(a) = (0, \ldots, 0, 1, 0, \ldots, 0)$ is a unit vector associated with the element $a \in \Omega_N$,

$$E(X_N^k) = \xi_0 \left(\prod_{t=1}^{N} M_t \right) V_k', \qquad k = 1, 2, \ldots, \tag{2.6}$$

where $V_k = \sum_{x=0}^{l} x^k U(C_x)$, and

$$\varphi_{X_N}(s) = \xi_0 \left(\prod_{t=1}^{N} M_t \right) W_s', \tag{2.7}$$

where $W_s = \sum_{x=0}^{l} e^{sx} U(C_x)$.

3. The main result

Our main goal in this section is to find the joint distribution of the random variables R_N (the number of rises) and D_N (the number of falls) in a random permutation $\pi(N) \in \mathcal{H}(S_N)$. To achieve this aim, we show that the random variable (or vector) $X_N = (R_N, D_N)$ is finite Markov chain imbeddable by constructing (i) a finite Markov chain $\{Y_t\}$ defined on a sequence of state spaces $\{\Omega_t\}$ with transition probability matrices $\{M_t\}$, and (ii) a partition $\{C_{(r,d)}^N\}$ on Ω_N such that

$$P(X_N = (r, d)) = P(Y_N \in C_{(r,d)}^N).$$

For $t = 1, \ldots, N$, we define the sequence of state spaces

$$\Omega_t = \{(R_t, D_t, L_t(k_t, \pi(t))) : \text{where } R_t, D_t \text{ and } L_t(k_t, \pi(t)) \text{ are defined in Section 2 and satisfy Lemmas 2.1 to 2.3}\}. \tag{3.1}$$

Similarly, we define the sequence of mappings (random vectors) $Y_t : \mathcal{H}(S_t) \to \Omega_t$ as

$$Y_t(\pi(t)) = (R_t, D_t, L_t(k_t, \pi(t))), \qquad \text{for all} \quad \pi(t) \in \mathcal{H}(S_t). \tag{3.2}$$

For $t = N$, $k_N = n$, and given (r, d) where r and d satisfy Lemma 2.1, we define the partition $\{C_{(r,d)}^N\}$ on the state space Ω_N as

$$C_{(r,d)}^N = \{(r, d, l) : l \text{ satisfies Lemmas 2.1 to 2.3}\}. \tag{3.3}$$

THEOREM 3.1. *Let $\mathcal{H}(S_N)$ be the collection of all possible random permutations generated by the integers in S_N with specification $[s] = [s_1, \ldots, s_n]$. Then, with respect to the random insertion procedure,*

(i) *the sequence of random vectors* $\{Y_t : \mathcal{H}(S_t) \to \Omega_t\}$ *defined by* (3.2) *form a finite Markov chain having transition probability matrices*

$$M_t = [p_{(r,d,l)(x,y,z)}(t)], \quad t = 1, 2, \ldots, N, \tag{3.4}$$

where, for every pair of $(r, d, l) \in \Omega_{t-1}$ *and* $(x, y, z) \in \Omega_t$, *the transition probabilities are defined as*

$$\begin{aligned} p_{(r,d,l)(x,y,z)}(t) &= P(Y_t = (x,y,z) \mid Y_{t-1} = (r,d,l)) \\ &= \begin{cases} I(t)(2s_{k_{t-1}}(t-1) - l)/t, & if \quad x = r, y = d, z = I(t)(l+1) \\ (r - I(t)(s_{k_{t-1}}(t-1) - l))/t, & if \quad x = r, y = d+1, z = I(t)l \\ (d - I(t)(s_{k_{t-1}}(t-1) - l))/t, & if \quad x = r+1, y = d, z = I(t)l \\ ((t - r - d) - I(t)l)/t, & if \quad x = r+1, y = d+1, z = I(t)l \\ 0, & otherwise, \end{cases} \end{aligned} \tag{3.5}$$

where the index functions $I(t)$ *are defined in* (2.4), *and*

(ii) *if* $\xi_0 = P(Y_0 = (0,0,0)) \equiv 1$ *is the initial probability, then the joint distribution of the random variables* R_N *and* D_N *is given by*

$$P((R_N, D_N) = (r,d)) = \xi_0 \left(\prod_{t=1}^{N} M_t \right) U'(C^N_{(r,d)}), \tag{3.6}$$

where $U'(C^N_{(r,d)})$ *is the transpose of* $U(C^N_{(r,d)})$,

$$U(C^N_{(r,d)}) = \sum_{(r,d,l) \in C^N_{(r,d)}} U((r,d,l)),$$

and $U((r,d,l)) = (0, \ldots, 0, 1, 0, \ldots, 0)$ *is a row unit vector with* 1 *at the coordinate associated with state* (r, d, l) *and* 0 *otherwise.*

PROOF. To prove the result (i), it is sufficient to prove that sequence $\{Y_t\}$, under the insertion process, has one-step transition matrices M_t with probabilities defined in (3.5). For $\pi(t-1) \in \mathcal{H}(S_{t-1})$, let $Y_{t-1}(\pi(t-1)) = (R_{t-1}, D_{t-1}, L_{t-1}(k_{t-1}, \pi(t-1)))$, where R_{t-1}, D_{t-1} and $L_{t-1}(k_{t-1}, \pi(t-1))$ are defined in Section 2 and satisfy Lemmas 2.1 to 2.3. Assume that the random permutation $\pi(t-1)$ has $R_{t-1} = r$ rises, $D_{t-1} = d$ falls, and $L_{t-1}(k_{t-1}, \pi(t-1)) = l$ levels associated with integer k_{t-1}. In the following, consider the transition probabilities when the integer k_t is inserted into one of the t gaps of the permutation $\pi(t-1)$. For $k_t = k_{t-1}$ (i.e. $I(t) = 1$), there are four possible cases:

(a) If k_t is inserted into one of the $(2s_{k_{t-1}}(t-1) - l)$ gaps of rises, falls, and levels associated with integer $k_t (= k_{t-1})$ (type **I** gaps) in the permutation $\pi(t-1)$, then $(r,d,l) \to (r,d,l+1)$ and the transition probability is

$$P(Y_t = (r,d,l+1) \mid Y_{t-1} = (r,d,l)) = (2s_{k_{t-1}}(t-1) - l)/t. \tag{3.7}$$

(b) If k_t is inserted into one of the $(r - (s_{k_{t-1}}(t-1) - l))$ gaps of rises not associated with the integer k_t (type **II** gaps), then (r,d,l) becomes $(r, d+1, l)$ and the transition probability is

$$P(Y_t = (r,d+1,l) \mid Y_{t-1} = (r,d,l)) = (r - (s_{k_{t-1}}(t-1) - l))/t. \tag{3.8}$$

(c) If k_t is inserted into one of the $(d-(s_{k_{t-1}}(t-1)-l))$ gaps of falls not associated with the integer k_t (type **III** gaps), then (r,d,l) becomes $(r+1,d,l)$ and the transition probability becomes

$$P(Y_t = (r+1,d,l) \mid Y_{t-1} = (r,d,l)) = (d-(s_{k_{t-1}}(t-1)-l))/t. \tag{3.9}$$

(d) If k_t is inserted into one of the $(t-r-d-l)$ gaps of levels not associated with the integer k_t (type **IV** gaps), then $(r,d,l) \to (r+1,d+1,l)$ and the transition probability is

$$P(Y_t = (r+1,d+1,l) \mid Y_{t-1} = (r,d,l)) = (t-r-d-l)/t. \tag{3.10}$$

Similarly, if $k_t = k_{t-1}+1$, then $I(t) = 0$ and the permutation $\pi(t-1)$ contains no integer k_t. Further, it follows that

$$R_{t-1}(k_t,\pi(t-1)) = D_{t-1}(k_t,\pi(t-1)) = L_t(k_t,\pi(t-1)) \equiv 0,$$

and the permutation $\pi(t)$ generated from $\pi(t-1)$ by insertion of integer k_t has no level associated with the integer k_t. Again, there are four cases for the transition probability of inserting k_t:

(a') Since there are no gaps of rises, falls and levels associated with integer k_t (type **I** gaps) in $\pi(t-1)$, then $(r,d,l) \to (r,d,0)$ and the transition probability is zero:

$$P(Y_t = (r,d,0) \mid Y_{t-1} = (r,d,l)) = 0. \tag{3.11}$$

(b') If k_t is inserted randomly into one of the r gaps of rises (type **II** gaps, no gaps of rises associated with integer k_t), then $(r,d,l) \to (r,d+1,0)$ and the transition probability becomes

$$P(Y_t = (r,d+1,0) \mid Y_{t-1} = (r,d,l)) = r/t. \tag{3.12}$$

(c') If k_t is inserted randomly into one of the d gaps (type **III** gaps, no gaps of falls associated with the integer k_t), then $(r,d,l) \to (r+1,d,0)$ and the transition probability is

$$P(Y_t = (r+1,d,0) \mid Y_{t-1} = (r,d,l)) = d/t. \tag{3.13}$$

(d') If k_t is inserted randomly into one of the $(t-r-d)$ gaps of levels (type **IV** gaps), then $(r,d,l) \to (r+1,d+1,0)$ and the transition probability becomes

$$P(Y_t = (r+1,d+1,0) \mid Y_{t-1} = (r,d,l)) = (t-r-d)/t. \tag{3.14}$$

Combining (3.7) with (3.11), (3.8) with (3.12), (3.9) with (3.13), and (3.10) with (3.14), yields (3.5). This proves that $\{Y_t\}$ forms a finite Markov chain with transition probability matrices defined by (3.5).

Given $X_N = (r,d)$, it follows from the definition of the partition $\{C^N_{(r,d)}\}$ on Ω_N that, for every (r,d),

$$P(X_N = (r,d)) = P(Y_N \in C^N_{(r,d)} \mid \xi_0).$$

Hence the random vector $X_N = (R_N, D_N)$ is finite Markov chain imbeddable. Result (ii) is a direct consequence of the Chapman-Kolmogorov equation in matrix form. □

Note that the above results proved by the insertion of integers following an increasing order of S_N can also be proved via an insertion of integers in S_N following a decreasing order.

It is easy to see from the above proof that, for every state $(r,d,l) \in \Omega_{t-1}$, the numbers of gaps of the four types (**I**, **II**, **III**, and **IV**) with respect to the inserting integer k_t are, respectively,

$$\begin{cases} g_1((r,d,l),k_t) = I(t)(2s_{k_{t-1}}(t-1)-l), \\ g_2((r,d,l),k_t) = (r - I(t)(s_{k_{t-1}}(t-1)-l)), \\ g_3((r,d,l),k_t) = (d - I(t)(s_{k_{t-1}}(t-1)-l)), \\ g_4((r,d,l),k_t) = ((t-r-d) - I(t)l). \end{cases} \tag{3.15}$$

The transition probabilities for each state $(r,d,l) \in \Omega_{t-1}$ can then be expressed as

$$(r,d,l) \to \begin{cases} (r,d,I(t)(l+1)), & \text{with probability } g_1/t \\ (r,d+1,I(t)l), & \text{with probability } g_2/t \\ (r+1,d,I(t)l), & \text{with probability } g_3/t \\ (r+1,d+1,I(t)l), & \text{with probability } g_4/t. \end{cases} \tag{3.16}$$

Equations (3.15) and (3.16) hence provide a simple sequential procedure to construct the state spaces $\{\Omega_t\}$,

$$\Omega_t = \cup_{(r,d,l)\in\Omega_{t-1}}\{(r,d,I(t)(l+1)),(r,d+1,I(t)l),$$
$$(r+1,d,I(t)l),(r+1,d+1,I(t)l)\},$$

and the transition probability matrices $\{M_t\}$. To illustrate this simple procedure, a detailed numerical example is given in Section 4.

4. Discussion and example

Given S_N, the sequence of integers $\{k_t\}$ is uniquely defined. The insertion procedure defined by the sequence of integers $\{k_t\}$ and the definitions of R_t, D_t, and $L_t(k_t,\pi(t))$, along with Lemmas 2.1 to 2.3 and Equations (3.15) and (3.16), provides a simple algorithm for constructing the state spaces $\{\Omega_t\}$ and the transition probability matrices $\{M_t\}$. To make our main results and their computational aspects more transparent, a detailed example is given below.

Example. Consider a set of integers $S_5 = \{1,1,2,3,3\}$ having specification $[s] = [2,1,2]$. The partially ordered sequence of integers $\{k_t; t = 1,\ldots,5\} = \{1,1,2,3,3\}$ is induced by S_5 and the index set $\{t = 1,2,3,4,5\}$. From Lemmas 2.1, 2.2, and 2.3, the insertion procedure, and (3.15) and (3.16), we obtain the state spaces $\Omega_t = \{(R_t,D_t,L_t(k_t,\pi(t)))\}$ and the transition probability matrices $M_t = [p_{(r,d,l),(x,y,z)}(t)]$ as follows:

$$\Omega_1 = \{(1,1,0)\}, \quad \Omega_2 = \{(1,1,1)\},$$
$$\Omega_3 = \{(1,2,0),(2,1,0),(2,2,0)\},$$
$$\Omega_4 = \{(1,3,0),(3,1,0),(2,2,0),(2,3,0),(3,2,0)\},$$
$$\Omega_5 = \{(1,3,1),(2,3,0),(2,4,0),(3,1,1),(3,2,0),$$
$$(4,2,0),(2,2,1),(3,3,0),(2,3,1),(3,2,1)\},$$

and

$$M_1 = M_2 = I_{1\times 1}, \qquad M_3 = [1/3, 1/3, 1/3]_{1\times 3},$$

$$M_4 = \begin{bmatrix} 1/4 & 0 & 1/2 & 1/4 & 0 \\ 0 & 1/4 & 1/2 & 0 & 1/4 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{bmatrix}_{3\times 5},$$

$$M_5 = \begin{bmatrix} 2/5 & 2/5 & 1/5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2/5 & 2/5 & 1/5 & 0 & 0 & 0 & 0 \\ 0 & 1/5 & 0 & 0 & 1/5 & 0 & 2/5 & 1/5 & 0 & 0 \\ 0 & 0 & 1/5 & 0 & 0 & 0 & 0 & 2/5 & 2/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/5 & 0 & 2/5 & 0 & 2/5 \end{bmatrix}_{5\times 10}.$$

The partition $\{C^N_{(r,d)}\}$ on the state space Ω_5 is then

$$C^5_{(3,1)} = \{(3,1,1)\}, \qquad C^5_{(1,3)} = \{(1,3,1)\}, \qquad C^5_{(2,2)} = \{(2,2,1)\},$$
$$C^5_{(3,2)} = \{(3,2,0),(3,2,1)\}, \qquad C^5_{(2,3)} = \{(2,3,0),(2,3,1)\}, \qquad C^5_{(3,3)} = \{(3,3,0)\},$$
$$C^5_{(4,2)} = \{(4,2,0)\}, \qquad C^5_{(2,4)} = \{(2,4,0)\}.$$

Take the initial probability $\xi_0 = 1$. It follows from (3.6) that, for all possible (r, d),

$$P((R_5, D_5) = (r,d)) = \xi_0 \left(\prod_{t=1}^{5} M_t\right) U'(C^5_{(r,d)}).$$

Hence we have

$$P((3,1)) = 1/30, \qquad P((1,3)) = 1/30, \qquad P((2,2)) = 2/15,$$
$$P((3,2)) = 1/5, \qquad P((2,3)) = 1/5, \qquad P((3,3)) = 4/15,$$
$$P((4,2)) = 1/15, \qquad P((2,4)) = 1/15.$$

The above results can be checked directly by writing out all possible $(5!/(2!2!) = 30)$ random permutations in $\mathcal{H}(S_5)$ according to the partition $\{C^N_{(r,d)}\}$ as follows:

$$C^5_{(3,1)} \sim \{\pi : (11233)\}, \qquad C^5_{(1,3)} \sim \{\pi : (33211)\},$$
$$C^5_{(2,2)} \sim \{\pi : (11332),(21133),(23311),(33112)\},$$
$$C^5_{(3,2)} \sim \{\pi : (11323),(13312),(31123),(12331),(12133),(23113)\},$$
$$C^5_{(2,3)} \sim \{\pi : (33121),(31132),(13321),(32113),(32311),(21331)\},$$
$$C^5_{(3,3)} \sim \{\pi : (13132),(13231),(13213),(31312),(31231),(31213),(21313),(23131)\},$$
$$C^5_{(4,2)} \sim \{\pi : (13123),(12313)\}, \qquad C^5_{(2,4)} \sim \{\pi : (31321),(32131)\}.$$

The moments and joint moments of R_5 and D_5 can then be calculated by a simple application of (2.6). For example, the means and expected product of R_5 and D_5 can be computed easily via $E(\cdot) = \xi_0(\prod_{t=1}^{5} M_t)V'$, where V equals (i) $V(rC^5_{(r,d)}) = (1,2,2,3,3,4,2,3,2,3)$, (ii) $V(dC^5_{(r,d)}) = (3,3,4,1,2,2,2,3,3,2)$, and (iii) $V(rdC^5_{(r,d)}) = (3,6,8,3,6,8,4,9,6,6)$, respectively. Numerically, this yields $E(R_5) = 2.6$, $E(D_5) = 2.6$, and $E(R_5D_5) = 6.6$. Similarly, the corresponding variances and covariance are $\mathrm{Var}(R_5) = 0.44$, $\mathrm{Var}(D_5) = 0.44$, and $\mathrm{Cov}(R_5, D_5) = -0.16$.

In the above example, the state spaces $\{\Omega_t\}$ and the transition probability matrices $\{M_t\}$ were constructed sequentially. To demonstrate this procedure, let us consider a set of integers $S_5^* = \{1,1,2,3,4\}$, which differs from the set $S_5 = \{1,1,2,3,3\}$ only by its last integer ("4" instead of "3"). The first four state spaces Ω_i^* and transition probability matrices M_i^*, $i = 1,2,3,4$, associated with S_5^* are the same as Ω_i and M_i for S_5, respectively. By (3.15) and (3.16), the state space Ω_5^* and transition probability matrix M_5^* can be constructed by inserting the integer "4" ($I(5) \equiv 0$) randomly into one of the four types of gaps for each state (r,d,l) in $\Omega_4^* = \{(1,3,0),(3,1,0),(2,2,0),(2,3,0),(3,2,0)\}$ in the following way:

(i) For state (1,3,0), the numbers of gaps of the four types with respect to the inserting integer "4" are $g_1 = 0$, $g_2 = 1$, $g_3 = 3$, and $g_4 = 1$, respectively. It follows that (1,3,0) goes to (1,4,0) with probability 1/5, to (2,3,0) with probability 3/5, and to (2,4,0) with probability 1/5.

(ii) For (3,1,0), the four numbers of gaps are $g_1 = 0$, $g_2 = 3$, $g_3 = 1$, $g_4 = 1$, and hence (3,1,0) goes to (3,2,0) with probability 3/5, to (4,1,0) with probability 1/5, and to (4,2,0) with probability 1/5.

(iii) For (2,2,0), it follows that $g_1 = 0$, $g_2 = 2$, $g_3 = 2$, $g_4 = 1$, and then (2,2,0) goes to (2,3,0) with probability 2/5, to (3,2,0) with probability 2/5, and to (3,3,0) with probability 1/5.

(iv) For (2,3,0), it follows that $g_1 = 0$, $g_2 = 2$, $g_3 = 3$, $g_4 = 0$, and then (2,3,0) goes to (2,4,0) with probability 2/5, and to (3,3,0) with probability 3/5.

(v) For (3,2,0), it follows that $g_1 = 0$, $g_2 = 3$, $g_3 = 2$, $g_4 = 0$, and hence (3,2,0) goes to (3,3,0) with probability 3/5, and to (4,2,0) with probability 2/5.

From (i) to (v), this yields the state space

$$\Omega_5^* = \{(1,4,0),(2,3,0),(2,4,0),(3,2,0),(4,1,0),(4,2,0),(3,3,0)\}$$

with transition probability matrix

$$M_5^* = \begin{array}{c} \\ (1,3,0) \\ (3,1,0) \\ (2,2,0) \\ (2,3,0) \\ (3,2,0) \end{array} \begin{array}{c} \begin{array}{ccccccc} (1,4,0) & (2,3,0) & (2,4,0) & (3,2,0) & (4,1,0) & (4,2,0) & (3,3,0) \end{array} \\ \left[\begin{array}{ccccccc} 1/5 & 3/5 & 1/5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3/5 & 1/5 & 1/5 & 0 \\ 0 & 2/5 & 0 & 2/5 & 0 & 0 & 1/5 \\ 0 & 0 & 2/5 & 0 & 0 & 0 & 3/5 \\ 0 & 0 & 0 & 0 & 0 & 2/5 & 3/5 \end{array}\right] \end{array}.$$

The aforementioned state space Ω_5 and transition matrix M_5 were obtained similarly by applying this procedure with insertion of the integer "3" ($I(5) \equiv 1$) for every state in the state space Ω_4. Simple computer programs for constructing the state spaces and transition matrices can be easily written, and we leave these details to the interested reader.

In view of the construction procedure for sequences $\{\Omega_t\}$ and $\{M_t\}$, it is clear that the joint distribution of R_N and D_N depends strongly on the structure of S_N. The number of states in the space Ω_N tends to infinity with an order less than $N^2 \max(s_1, \ldots, s_n)$, and comparing this with the number of permutations $N!/\prod_{i=1}^n (s_i!)$ in $\mathcal{H}(S_N)$, which tends to infinity exponentially fast, our proposed approach is more efficient computationally.

The main intuitive reason why our results hold can be summarized briefly as follows. The random vector $Y_t = (R_t, D_t, L_t(k_t, \pi(t)))$ contains sufficient information to guarantee that the sequence $\{Y_t\}$ forms a Markov chain with respect to the insertion procedure. Hence the joint distribution of R_N and D_N is a direct consequence of the projection $(R_t, D_t, L_t) \to (R_t, D_t)$ and the Chapman-Kolmogorov equation.

References

Carlitz, L. (1964). Extended Bernoulli and Eulerian numbers, *Duke Math. J.*, **31**, 667–690.

Carlitz, L. (1972). Enumeration of sequences by rises and falls: A refinement of the Simon Newcomb problem, *Duke Math. J.*, **39**, 267–280.

Carlitz, L. (1974). Permutations and sequences, *Adv. Math.*, **14**, 92–120.

David, F. N. and Barton, D. E. (1962). *Combinatorial Chance*, Hafiner, New York.

Dillon, J. F. and Roselle, D. (1969). Simon Newcomb's problem, *SIAM J. Appl. Math.*, **17**, 1086–1093.

Fu, J. C. (1995). Exact and limiting distribution of the number of successions in a random permutations, *Ann. Inst. Statist. Math.*, **47**, 435–446.

Fu, J. C. and Koutras, M. V. (1994). Distribution theory of runs: a Markov chain approach, *J. Amer. Statist. Assoc.*, **89**, 1050–1058.

Fu, J. C., Lou, W. Y. W. and Wang, Y. J. (1998). On exact distributions of Eulerian and Simon Newcomb numbers associated with random permutations, *Statist. Probab. Lett.*, **42**, 115–125.

Giladi, E. and Keller, J. B. (1994). Eulerian number asymptotics, *Proc. Roy. Soc. London Ser. A*, **445**, 291–303.

Harris, B. and Park, C. J. (1994). A generalization of the Eulerian numbers with a probabilistic application, *Statist. Probab. Lett.*, **20**, 37–47.

MacMahon, P. A. (1915). *Combinatory Analysis*, Cambridge, London.

Nicolas, J. L. (1992). An integral representation for Eulerian numbers, *Colloq. Math. Soc. János Bolyai*, **60**, 513–527.

Riordan, J. (1958). *An Introduction to Combinatorial Analysis*, Wiley, New York.

Takacs, L. (1979). A generalization of the Eulerian numbers, *Publ. Math. Debrecen*, **26**, 173–181.

Tanny, S. (1973). A probabilistic interpretation of Eulerian numbers, *Duke Math. J.*, **40**, 717–722.