Ann. Inst. Statist. Math.
Vol. 51, No. 1, 1–15 (1999)

DISTRIBUTIONS OF RUNS AND CONSECUTIVE SYSTEMS ON DIRECTED TREES*

SIGEO AKI

Department of Mathematical Science, Osaka University, Toyonaka, Osaka 560-8531, Japan

(Received March 27, 1997; revised November 10, 1997)

Abstract. In this paper we study exact distributions of runs on directed trees. On the assumption that the collection of random variables indexed by the vertices of a directed tree has a directed Markov distribution, the exact distribution theory of runs is extended from based on random sequences to based on directed trees. The distribution of the number of success runs of a specified length on a directed tree along the direction is derived. A consecutive-k-out-of-n:F system on a directed tree is introduced and investigated. By assuming that the lifetimes of the components are independent and identically distributed, we give the exact distribution of the lifetime of the consecutive system. The results are not only theoretical but also suitable for computation.

Key words and phrases: Probability generating function, discrete distributions, run, directed tree, graph, reliability, Markov tree, lifetime, consecutive system, order statistics.

1. Introduction and preliminaries

The exact distribution theory of runs of a specified length in random sequences has been developed by many authors (cf. e.g. Rajarshi (1974), Schwager (1983), Philippou and Muwafi (1982), Philippou *et al.* (1983), Aki (1985), Philippou (1986), Hirano (1986), Philippou and Makri (1986), Ebneshahrashoob and Sobel (1990), Aki (1992), Hirano and Aki (1993), Balasubramanian *et al.* (1993), Godbole and Papastavridis (1994), Mohanty (1994), Fu and Koutras (1994), Aki and Hirano (1995), Koutras and Alexandrou (1995), Uchida and Aki (1995), Aki *et al.* (1996), Balakrishnan et al. (1997) and the references therein).

The traditional approach for the problem is based on enumerative combinatorics. As order of dependency of underlying random sequences grows higher, however, new powerful methods such as the method of conditional probability generating functions (pgf's) and the method of Markov chain imbedding have been developed. The method of conditional pgf's was introduced in the problem by Ebneshahrashoob and Sobel (1990) and used effectively by Aki (1992), Uchida

* This research was partially supported by the ISM Cooperative Research Program (96-ISM-CRP-11).

and Aki (1995), Aki *et al.* (1996) and others. The method of Markov chain imbedding was introduced by Fu and Koutras (1994) and developed by Koutras and Alexandrou (1995), Fu (1996) and Lou (1996). The method of conditional *pgf*'s is more general than the method of Markov chain imbedding. However, when the method of conditional *pgf*'s is applied to derivation of exact distributions of runs in random sequences, the resulting system of equations of conditional *pgf*'s often becomes linear even if order of dependency of the random sequences is very high (cf. Aki *et al.* (1996)). The linearlity of the system enables us to consider a virtual Markov chain and to obtain fruitful results from the setup.

Among various applications of the theory, in particular, consecutive engineering systems such as linear consecutive-k-out-of-n:F systems and m-consecutive-k-out-of-n:F systems are very interesting and stimulative (cf. e.g. Hirano (1994) and Chao *et al.* (1995)). The reliability and the lifetime of each of the systems are directly related to the distribution of the number of runs of length k in the corresponding sequence of random variables.

In this paper we consider exact distributions of runs on directed trees. Typical practical examples of consecutive systems are a sequence of n microwave stations which transmits information from place A to place B and a sequence of n pump stations for transporting oil by pipes from point A to point B and so on. It is much more practical to assume that the components (microwave stations or pump stations) are not necessarily placed in line or that the number of destinations is not necessarily one, say, from A to B, C and D. The most simplest statistical model for the problem is a collection of random variables indexed by the vertices of a directed tree. Figure 1 shows an example of a realization of a directed tree of binary ({ •, ∘ }-valued) random variables.

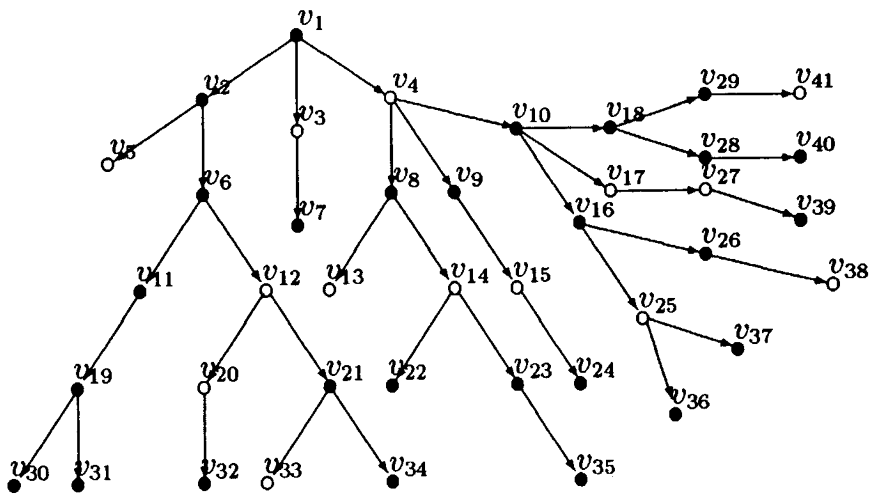

Fig. 1. A realization of a collection of binary random variables indexed by the vertex set of a directed tree.

For basic notions in graphical models we refer the reader to Lauritzen (1996) and Ripley (1996). If every vertex except for a leaf has only one child, then the directed tree reduces to a sequence of finite length. Therefore, our model includes usual linear consecutive systems. By assuming that the collection of random variables indexed by the vertices of a directed tree has a directed Markov distribution (Lauritzen ((1996), p. 52)), we can extend the exact distribution theory of runs from based on sequences to based on directed trees. For the purpose the method of conditional *pgf*'s is still a powerful tool. However, if the directed tree is not a sequence, the system of equations of conditional *pgf*'s is no longer linear (see Theorem 2.1) and the method of Markov chain imbedding may not be available.

Here, we briefly explain about the Markov properties on graphs for the readers' convenience. Some different Markov properties such as the pairwise, the local, and the global Markov properties on graphs have been introduced and investigated by many researchers (cf. e.g. Lauritzen (1996)). The followings are well known on directed acyclic graphs. Let P be a probability distribution of a collection of random variables indexed by the vertices of a directed acyclic graph $\mathcal{G}$ and let V be the totality of the vertices of $\mathcal{G}$. For $A \subset V$, we denote $X_A \equiv \{X_a, a \in V\}$. We say that P obeys the directed global Markov property relative to $\mathcal{G}$ if X_A and X_B are conditionally independent given X_S whenever A and B are separated by S in the moral graph of the smallest ancestral set containing $A \cup B \cup S$. Here, the moral graph of a directed acyclic graph $\mathcal{G}$ is the undirected graph with the same vertex set obtained by adding an edge between parents when they have a common child and by ignoring the direction. In particular, in case of directed trees treated in this paper the moral graph of a directed tree is the undirected graph obtained only by ignoring the direction, since every vertex does not have more than or equal to two parents in any directed tree. We say that P obeys the directed local Markov property if any random variable X_a is conditionally independent of its non-descendants given its parents. Then, P obeys the directed global Markov property relative to $\mathcal{G}$ if and only if P obeys the directed local Markov property relative to $\mathcal{G}$ (Theorem 3.27 of Lauritzen (1996)). From the equivalence of the Markov properties on directed acyclic graphs, P is said to have a directed Markov distribution if P obeys either of the Markov properties. In this paper we restrict ourselves to study the probability distributions on directed trees which are the simplest special graphs of directed acyclic graphs.

We give in Section 2 a general result for obtaining the exact distribution of number of runs on directed trees by extending the method of conditional *pgf*'s. The result provides a feasible algorithm for the distribution of runs (or reliability of consecutive systems on directed trees).

If lifetimes of the components of a linear consecutive-k-out-of-n:F system are independent and identically distributed, the distribution of the lifetime of the system can be written as a finite mixture of distributions of n order statistics of the lifetimes of the components (Aki and Hirano (1996)); the mixing weights, which do not depend on the distribution of the lifetime of each component, are obtained from the numbers of minimal m-cutsequences for $m = 1, 2, \ldots, n$. The result is very useful for obtaining moments of the lifetime of the system, since the moments of order statistics have been studied very well for common lifetime distributions. In

Section 3, we give a method for deriving the numbers of minimal m-cutsequences of consecutive systems on directed trees and obtain the distribution and some characteristics of the lifetime of the consecutive systems on directed trees.

Before stating the general result, we illustrate how to enumerate the number of runs on directed trees by using Fig. 1. We can find out the following "$\bullet$"-runs of length 3 along the direction on the directed tree in Fig. 1: $R_1 = (v_1, v_2, v_6)$, $R_2 = (v_{11}, v_{19}, v_{30})$, $R_3 = (v_{11}, v_{19}, v_{31})$, $R_4 = (v_{10}, v_{16}, v_{26})$, $R_5 = (v_{10}, v_{18}, v_{28})$ and $R_6 = (v_{10}, v_{18}, v_{29})$. Here, we adopt the non-overlapping counting method along the direction in every path from the root v_1 to one of the leaves $\{ v_5, v_{30}, v_{31}, v_{32}, v_{33}, v_{34}, v_7, v_{13}, v_{22}, v_{35}, v_{24}, v_{36}, v_{37}, v_{38}, v_{39}, v_{40}, v_{41} \}$. Here, we do not count the runs (v_2, v_6, v_{11}) or (v_6, v_{11}, v_{19}) since they overlap with the run $R_1 = (v_1, v_2, v_6)$ in the path $v_1 \to v_2 \to v_6 \to v_{11} \to v_{19} \to v_{30}$. Though R_2 overlaps with R_3, we count the both runs since they can not be included simultaneously in a path from the root to a leaf. We do not count the run (v_{28}, v_{18}, v_{29}) because the run is not along the direction.

2. Number of runs on directed trees

Let T be a directed tree and let V be the totality of the vertices of T. Suppose we are given a collection of $\{0,1\}$-valued random variables $\{X_v, v \in V\}$. Let v_0 be the root of T. Of course, we assume that all edges are directed away from the root. Further, we assume that $\{X_v, v \in V\}$ has the directed Markov distribution (cf. Lauritzen ((1996), p. 52) or Ripley ((1996), p. 253)), with the initial distribution at the root $P(X_{v_0} = 1) = p = 1 - q$ and the conditional probabilities

$$P(X_v = 1 | X_{pa(v)} = 1) = p_1 = 1 - q_1,$$
$$P(X_v = 1 | X_{pa(v)} = 0) = p_0 = 1 - q_0,$$

for each vertex v except for the root, where $pa(v)$ denotes the parent of the vertex v. Then, the collection of the random variables $\{X_v, v \in V\}$ is often called a Markov tree. We fix any vertex v except for the root. Suppose that the vertex has $a(v)$ ancestors $v^1, v^2, \dots, v^{a(v)}$ with $pa(v^j) = v^{j+1}$ for $j = 1, 2, \dots, a(v) - 1$. Assume also that the vertex v has $c(v)$ children $v_1, v_2, \dots, v_{c(v)}$. We denote by T_v the subtree which consists of the vertex v (the root of the subtree) and of all of the descendants of v. V_v denotes the set of vertices of T_v. Let $\phi_{v_0}(t)$ be the pgf of the distribution of the number of non-overlapping "1"-runs of length k along the direction in $\{X_v, v \in V\}$. For every vertex v except for the root v_0, we let $\phi_v^{(0)}(t)$ be the pgf of the conditional distribution of the number of non-overlapping "1"-runs of length k along the direction in $\{X_w, w \in V_v\}$ given that $X_{pa(v)} = 0$ and let $\psi_v^{(0)}(t)$ be the pgf of the conditional distribution of the number of non-overlapping "1"-runs of length k along the direction in $\{X_w, w \in V_v\}$ given that at the vertex $pa(v)$ a "1"-run of length k along the direction is observed. For $l = 1, 2, \dots, \min\{(k-1), a(v)\}$, we let $\phi_v^{(l)}(t)$ be the pgf of the conditional distribution of the number of non-overlapping "1"-runs of length k along the direction in $\{X_w, w \in V_v\}$ given that at the vertex $pa(v)$ a "1"-run of length l along the direction is observed. Then we have

THEOREM 2.1. *The pgf's $\phi_{v_0}(t)$, $\phi_v^{(l)}(t)$ and $\psi_v^{(0)}(t)$ for $l = 0, 1, \ldots, k-1$ satisfy the following recurrence relations*;

$$\text{(2.1)}\quad \phi_{v_0}(t) = \begin{cases} q\prod_{j=1}^{c(v_0)} \phi_{v_{0j}}^{(0)}(t) + p\prod_{j=1}^{c(v_0)} \phi_{v_{0j}}^{(1)}(t) & \text{if } c(v_0) > 0 \text{ and } k > 1 \\ q\prod_{j=1}^{c(v_0)} \phi_{v_{0j}}^{(0)}(t) + pt\prod_{j=1}^{c(v_0)} \psi_{v_{0j}}^{(0)}(t) & \text{if } c(v_0) > 0 \text{ and } k = 1 \\ q + pt & \text{if } c(v_0) = 0 \text{ and } k = 1 \\ 1 & \text{if } c(v_0) = 0 \text{ and } k > 1, \end{cases}$$

for $v \neq v_0$,

$$\text{(2.2)}\quad \phi_v^{(0)}(t) = \begin{cases} q_0\prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_0\prod_{j=1}^{c(v)} \phi_{v_j}^{(1)}(t) & \text{if } c(v) > 0 \text{ and } k > 1 \\ q_0\prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_0t\prod_{j=1}^{c(v)} \psi_{v_j}^{(0)}(t) & \text{if } c(v) > 0 \text{ and } k = 1 \\ q_0 + p_0t & \text{if } c(v) = 0 \text{ and } k = 1 \\ 1 & \text{if } c(v) = 0 \text{ and } k > 1, \end{cases}$$

and

$$\text{(2.3)}\quad \psi_v^{(0)}(t) = \begin{cases} q_1\prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_1\prod_{j=1}^{c(v)} \phi_{v_j}^{(1)}(t) & \text{if } c(v) > 0 \text{ and } k > 1 \\ q_1\prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_1t\prod_{j=1}^{c(v)} \psi_{v_j}^{(0)}(t) & \text{if } c(v) > 0 \text{ and } k = 1 \\ q_1 + p_1t & \text{if } c(v) = 0 \text{ and } k = 1 \\ 1 & \text{if } c(v) = 0 \text{ and } k > 1, \end{cases}$$

for $1 \leq l < \min\{(k-1), a(v)\}$ (automatically $k > 2$),

$$\text{(2.4)}\quad \phi_v^{(l)}(t) = \begin{cases} q_1\prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_1\prod_{j=1}^{c(v)} \phi_{v_j}^{(l+1)}(t) & \text{if } c(v) > 0 \\ 1 & \text{if } c(v) = 0, \end{cases}$$

for $k > 1$,

$$\phi_v^{(k-1)}(t) = \begin{cases} q_1 \prod_{j=1}^{c(v)} \phi_{v_j}^{(0)}(t) + p_1 t \prod_{j=1}^{c(v)} \psi_{v_j}^{(0)}(t) & \textit{if} \quad c(v) > 0 \\ q_1 + p_1 t & \textit{if} \quad c(v) = 0. \end{cases} \tag{2.5}$$

PROOF. Note that $\{X_v, v \in V\}$ follows a directed Markov distribution. Since every vertex does not have more than or equal to two parents in any directed tree, its moral graph is obtained only by ignoring the direction. Because the subsets $V_{v_1}, \dots, V_{v_{c(v)}}$ are separated by the parent $\{v\}$ in the moral graph of the smallest ancestral set containing V_v, from the directed global Markov property, $\{X_w, w \in V_{v_1}\}, \dots,$ and $\{X_w, w \in V_{v_{c(v)}}\}$ become conditionally independent when X_v is given. Then, since the number of "1"-runs observed in $\{X_w, w \in V_v\}$ is the sum of the (conditionally independent) numbers of "1"-runs observed in $\{X_w, w \in V_{v_1}\}$, ..., and $\{X_w, w \in V_{v_{c(v)}}\}$, the corresponding *pgf* becomes the product of the *pgf*'s relative to $\{X_w, w \in V_{v_j}\}$, $j = 1, \dots, c(v)$. Thus, we have (2.1), (2.2) and (2.3). Similarly, when $\{X_{v^1} = 1, \dots, X_{v^l} = 1\}$ is given, the numbers of "1"-runs observed in $\{X_w, w \in V_{v_j}\}$, $j = 1, \dots, c(v)$ become conditionally independent if X_v is also given and hence we have (2.4) and (2.5). This completes the proof.

Remark 2.1. In fact, the boundary conditions of the recurrence relations in Theorem 2.1 are given at every leaf since $c(v) = 0$ is observed at every leaf. However, by taking account of the length of the remaining subtree at every vertex, we can add another type of boundary conditions; denoting by $|T_v|$ the length of T_v (the maximum of the length of paths in T_v),

$$\begin{cases} \phi_v^{(l)}(t) = 1 & \text{if} \quad l + |T_v| + 1 < k \\ \psi_v^{(0)}(t) = 1 & \text{if} \quad |T_v| + 1 < k. \end{cases}$$

Remark 2.2. Letting $k = 1$ in Theorem 2.1, we obtain the distribution of the number of "1" in the Markov tree. If we extend the r-out-of-n:F system from usual i.i.d. case to the case of a Markov tree, we find out the reliability of the system immediately by calculating the probability that the number of "1" on the tree is less than r. In particular, when $k = 1$ and $p = p_0 = p_1$ (i.i.d. case), we see that $\phi_{v_0}(t) = (pt + q)^N$, where N is the cardinality of V, since $\phi_v^{(0)}(t) = \psi_v^{(0)}(t)$ holds for every $v \in V$. This is the usual binomial distribution.

Remark 2.3. When $c(v) = 0$ or 1 for every $v \in V$, the directed tree reduces to a sequence of a finite length and hence we can arrange the vertices in line as $v_0, v_1, \dots, v_{N-1}$, where N is the cardinality of V. Then the corresponding sequence $X_0, X_1, \dots, X_{N-1}$ becomes a homogeneous Markov chain. Therefore, we can reduce the set of the *pgf*'s by modifying the indices from every vertex to the length of the remaining sequence. For example, we substitute $\phi_{v_0}(t)$, $\phi_{v_j}^{(l)}(t)$ and $\psi_{v_j}^{(0)}(t)$ for $\phi_N(t)$, $\phi_{N-j}^{(l)}(t)$ and $\psi_{N-j}^{(0)}(t)$, respectively. Then, Theorem 2.1 and

Remark 2.1 imply the following linear system of equations;

$$\begin{cases} \phi_n^{(0)}(t) = 1 & \text{if} \quad 0 \le n < k \\ \psi_n^{(0)}(t) = 1 & \text{if} \quad 0 \le n < k \\ \phi_n^{(0)}(t) = q_0\phi_{n-1}^{(0)}(t) + p_0q_1\phi_{n-2}^{(0)}(t)\cdots & \\ \qquad +p_0p_1^{k-2}q_1\phi_{n-k}^{(0)}(t) + p_0p_1^{k-1}t\psi_{n-k}^{(0)}(t) & \text{if} \quad n \ge k \\ \psi_n^{(0)}(t) = q_1\phi_{n-1}^{(0)}(t) + p_1q_1\phi_{n-2}^{(0)}(t)\cdots & \\ \qquad +p_1^{k-1}q_1\phi_{n-k}^{(0)}(t) + p_1^k t\psi_{n-k}^{(0)}(t) & \text{if} \quad n \ge k. \end{cases}$$

This can be solved explicitly similarly as Theorem 3 of Aki and Hirano (1993).

When a collection of {0, 1}-valued random variables indexed by the vertices of a directed tree with a directed Markov distribution is given, Theorem 2.1 indeed provides an algorithm for deriving the *pgf* of the distribution of the number of "1"-runs of a specified length. We illustrate here how to derive the *pgf* by using the following example.

Example 2.1. We calculate the *pgf* of the distribution of the number of "1"-runs of length 3 on the directed tree in Fig. 2. For simplicity, we assume that $p = p_0 = p_1$, that is, $X_{v_j}, j = 1, 2, \ldots, 18$ are independent and identically distributed. Since the root v_1 has the children $\{v_2, v_3, v_4\}$, we have from (2.1)

$$\phi_{v_1}(t) = q\phi_{v_2}^{(0)}(t)\phi_{v_3}^{(0)}(t)\phi_{v_4}^{(0)}(t) + p\phi_{v_2}^{(1)}(t)\phi_{v_3}^{(1)}(t)\phi_{v_4}^{(1)}(t).$$

Next, noting that the vertex v_2 has children $\{v_5, v_6\}$, we have

$$\phi_{v_2}^{(0)}(t) = q\phi_{v_5}^{(0)}(t)\phi_{v_6}^{(0)}(t) + p\phi_{v_5}^{(1)}(t)\phi_{v_6}^{(1)}(t).$$

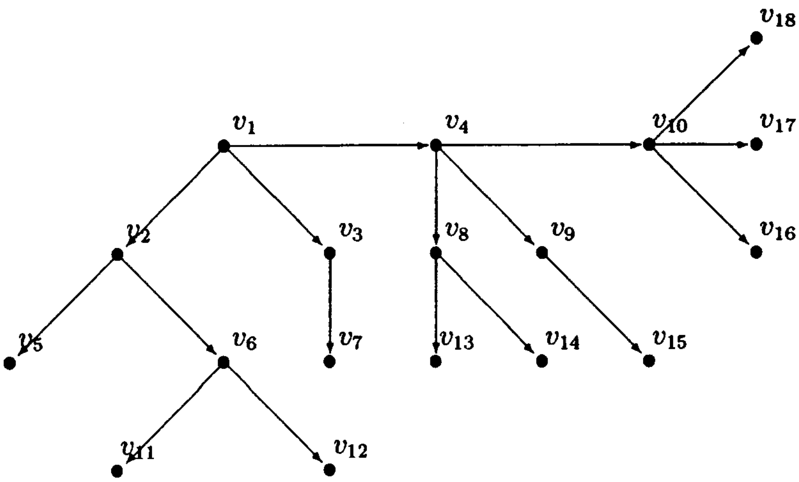

Fig. 2. An example of a directed tree.

Since the vertex v_5 does not have a child, we have $\phi_{v_5}^{(0)}(t) = 1$ and $\phi_{v_5}^{(1)}(t) = 1$ from (2.2) and (2.4), respectively. From Remark 2.1, we observe $\phi_{v_6}^{(0)}(t) = 1$. Further, by using (2.4), we have

$$\phi_{v_6}^{(1)}(t) = q + p\phi_{v_{11}}^{(2)}(t)\phi_{v_{12}}^{(2)}(t).$$

Here, from (2.5), we have that $\phi_{v_{11}}^{(2)}(t) = q+pt$ and $\phi_{v_{12}}^{(2)}(t) = q+pt$. Consequently, we obtain

$$\phi_{v_2}^{(0)}(t) = q + p(q + p(q + pt)^2).$$

Similarly, we have

$$\begin{aligned}
\phi_{v_3}^{(0)}(t) &= 1,\\
\phi_{v_4}^{(0)}(t) &= q + p(q + p(q + pt)^2)(q + p(q + pt))(q + p(q + pt)^3),\\
\phi_{v_2}^{(1)}(t) &= q + p(q + pt)^2,\\
\phi_{v_3}^{(1)}(t) &= q + p(q + pt),\\
\phi_{v_4}^{(1)}(t) &= q + p(q + pt)^3.
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\phi_{v_1}(t) = {} & q[q + p(q + p(q + pt)^2)]\\
& \times[q + p(q + p(q + pt)^2)(q + p(q + pt))(q + p(q + pt)^3)]\\
& + p[q + p(q + pt)^2][q + p(q + pt)][q + p(q + pt)^3].
\end{aligned}$$

Since the pgf is a polynomial with respect to t, we only take out the coefficients of t^j for deriving the probability distribution. In fact, we have

$$\begin{aligned}
\Pr(0) &= -p^{15} + 8p^{14} - 25p^{13} + 34p^{12} - 4p^{11} - 42p^{10} + 34p^9 + 14p^8 - 13p^7\\
&\quad -24p^6 + 15p^5 + 17p^4 - 14p^3 + 1\\
\Pr(1) &= (8p^{10} - 41p^9 + 66p^8 - 6p^7 - 77p^6 + 35p^5 + 43p^4 - 5p^3 - 38p^2 + 2p + 14)\\
&\quad \cdot(p-1)^2p^3\\
\Pr(2) &= -(28p^9 - 119p^8 + 145p^7 + 25p^6 - 138p^5 + 9p^4 + 58p^3 + 25p^2 - 29p - 9)\\
&\quad \cdot(p-1)^2p^4\\
\Pr(3) &= (56p^9 - 245p^8 + 345p^7 - 80p^6 - 182p^5 + 68p^4 + 58p^3 + 8p^2 - 27p - 2)\\
&\quad \cdot(p-1)p^5\\
\Pr(4) &= -(70p^7 - 245p^6 + 250p^5 - 98p^3 - 8p^2 + 20p + 12)(p-1)p^7\\
\Pr(5) &= (56p^6 - 147p^5 + 93p^4 + 24p^3 - 19p^2 - 12p - 2)(p-1)p^8\\
\Pr(6) &= -(28p^5 - 77p^4 + 61p^3 - 4p^2 - 7p - 2)p^{10}\\
\Pr(7) &= (8p+1)(p-1)^2p^{12}\\
\Pr(8) &= -(p-1)p^{14}.
\end{aligned}$$

As a special case of directed trees, we consider a complete m-ary tree of length $(n-1)$. We write the set of the vertices

$$V = \{v^{(0)}; v_0^{(1)}, v_1^{(1)}, \dots, v_{m-1}^{(1)}; v_{0,0}^{(2)}, \dots, v_{m-1,m-1}^{(2)}; \dots; v_{0,\dots,0}^{(n-1)}, \dots, v_{m-1,\dots,m-1}^{(n-1)}\},$$

where we assume that every vertex $v_{k_1,\dots,k_j}^{(j)}$ has just m children $\{v_{k_1,\dots,k_j,0}^{(j+1)}, \dots, v_{k_1,\dots,k_j,m-1}^{(j+1)}\}$ for $j = 0, 1, \dots, n-2$. For simplicity we assume that $\{X_v, v \in V\}$ is independent and identically distributed with $p = p_0 = p_1$. For v_1 and $v_2 \in V$ and for $l = 0, 1, \dots, k-1$, $\phi_{v_1}^{(l)}(t) = \phi_{v_2}^{(l)}(t)$ holds if $|T_{v_1}| = |T_{v_2}|$, since every vertex has just m children and $\{X_v, v \in V\}$ is independent and identically distributed in this model. We note also that $\phi_v^{(0)}(t) = \psi_v^{(0)}(t)$ for every $v \in V$. Hence, we denote for $l = 0, 1, \dots, k-1$, $\phi_{j+1}^{(l)}(t) = \phi_v^{(l)}(t)$ when $|T_v| = j$. Then we have from Theorem 2.1

COROLLARY 2.1. *The pgf's of the conditional distributions of the number of "1"-runs of length k on the i.i.d. complete m-ary tree of length $(n-1)$ satisfy the recurrence relations;*

$$\begin{cases} \phi_n(t) = q(\phi_{n-1}(t))^m + p(\phi_{n-1}^{(1)}(t))^m \\ \phi_{n-1}^{(1)}(t) = q(\phi_{n-2}(t))^m + p(\phi_{n-2}^{(2)}(t))^m \\ \dots \\ \phi_{n-k+1}^{(k-1)}(t) = q(\phi_{n-k}(t))^m + pt(\phi_{n-k}(t))^m, \end{cases}$$

with boundary conditions $\phi_j^{(l)}(t) = 1$ for $l + j < k$.

Remark 2.4. By assuming $m = 1$ in Corollary 2.1, we have the *pgf* of the distribution of the number of "1"-runs in i.i.d. sequence of length n, where the distribution is called the binomial distribution of order k (cf. Johnson *et al.* (1992)). When $m = 1$, the above system of equations is linear and we can reduce it to

$$\begin{cases} \phi_n(t) = \sum_{j=0}^{k-1} p^j q \phi_{n-j-1}(t) + p^k t \phi_{n-k}(t) & \text{if} \quad n \geq k \\ \phi_n(t) = 1 & \text{if} \quad n < k. \end{cases}$$

This was solved by Aki and Hirano (1988, Theorem 2.1) and $\phi_n(t)$ is written as

$$\sum_{m=0}^{k-1} \sum_{n_1+2n_2+\cdots+kn_k=n-m} \binom{n_1+n_2+\cdots+n_k}{n_1, n_2, \dots, n_k} p^n \left(\frac{q}{p}\right)^{n_1+\cdots+n_k} \left(1+\frac{p}{q}t\right)^{n_k}.$$

3. Lifetime of consecutive systems

In this section, we consider the distribution of the lifetime of a consecutive system every component of which is allocated on the corresponding vertex of a directed tree. Let T be a directed tree with the set of vertices V. Let N be the cardinality of V. Suppose that N components are allocated at the vertices V one by one. We assume that the system fails if and only if consecutive k components along the direction fail. We shall call it a consecutive-k-out-of-N:F system on T. A subset W $(\subset V)$ is called i-cutset if and only if W contains just i elements and the system fails if the corresponding components of the subset fail. A subsequence of elements of V of length i is called i-cutsequence if it is a permutation of an i-cutset. An i-cutsequence $(v_1, v_2, \dots, v_i)$ is said to be minimal if the subsequence $(v_1, v_2, \dots, v_{i-1})$ is not an $(i-1)$-cutsequence. Let $\xi_1, \dots, \xi_N$ be the lifetimes of the components. We assume that $\xi_1, \dots, \xi_N$ are independent and identically distributed with a cumulative distribution function $G(t)$. Let $\xi_{(1)} \le \cdots \le \xi_{(N)}$ be the order statistics of $\xi_1, \dots, \xi_N$ and let $G_{(i)}(t)$ be the cumulative distribution function of $\xi_{(i)}$. Then, from Aki and Hirano (1996), there exist constants $\omega_1, \dots, \omega_N$ ($\omega_i \ge 0$ for $i = 1, \dots, N$ and $\omega_1 + \cdots + \omega_N = 1$) such that the distribution function $F(t)$ of the lifetime of the system can be written as $F(t) = \sum_{i=1}^N \omega_i G_{(i)}(t)$. Here, ω_i is given explicitly as $\omega_i = r_{i,k}(N-i)!/N!$, where $r_{i,k}$ is the number of minimal i-cutsequences of the system.

Consequently, the problem we have to study is to find out the number of minimal i-cutsequences of the system. Let $b_{i,k}$ be the number of subsets B of V such that $|B| = i$ and B is not a cutset of the system, where $|B|$ denotes the cardinality of B. Then, it is easy to see that the number of minimal i-cutsequences can be written as

$$r_{i,k} = (N-i+1)((i-1)!)b_{i-1,k} - i!b_{i,k}.$$

We give how to calculate $b_{i,k}$ for $i = 0, 1, \dots, N$. We fix any vertex v. Suppose that the vertex v has $a(v)$ ancesters $v^1, v^2, \dots, v^{a(v)}$ with $pa(v^j) = v^{j+1}$ for $j = 1, 2, \dots, a(v) - 1$. Assume also that the vertex v has $c(v)$ children $v_1, v_2, \dots, v_{c(v)}$. For $i = 1, 2, \dots, \min\{(k-1), a(v)\}$, we let $P_i \equiv \{v^1, v^2, \dots, v^i\}$ and let $P_0 \equiv \phi$. We define

$$b_{v,m}^{(i)} \equiv \text{ number of } \{M \subset V_v \mid |M| = m \text{ and } M \cup P_i \text{ is not a cutset of the system}\}.$$

We write the generating function of $b_{v,m}^{(i)}$ as $\eta_v^{(i)}(t) = \sum_{m=0}^{|V_v|} b_{v,m}^{(i)} t^m$. Then, we have

THEOREM 3.1. *The generating functions $\eta_v^{(i)}(t)$ for $i = 0, 1, \dots, k-1$ satisfy the recurrence relations;*

$$\eta_v^{(i)}(t) = \begin{cases} \eta_{v_1}^{(0)}(t)\eta_{v_2}^{(0)}(t)\cdots\eta_{v_{c(v)}}^{(0)}(t) + t\eta_{v_1}^{(i+1)}(t)\eta_{v_2}^{(i+1)}(t)\cdots\eta_{v_{c(v)}}^{(i+1)}(t) \\ \qquad\qquad\qquad\qquad\qquad \text{if} \quad c(v) > 0 \quad \text{and} \quad 0 \le i < k-1 \\ (t+1) \qquad\qquad\qquad\quad\; \text{if} \quad c(v) = 0 \quad \text{and} \quad 0 \le i < k-1, \end{cases}$$

(3.1)

$$\eta_v^{(k-1)}(t) = \begin{cases} \eta_{v_1}^{(0)}(t)\eta_{v_2}^{(0)}(t)\cdots\eta_{v_{c(v)}}^{(0)}(t) & \text{if} \quad c(v) > 0 \\ 1 & \text{if} \quad c(v) = 0. \end{cases} \tag{3.2}$$

PROOF. Let M be a subset of V_v such that $M \cup P_i$ is not a cutset of the system. When $v \notin M$, then $M \cap V_{v_j}$ is not a cutset for every $j = 1, \dots, c(v)$ and M is written as $M = \cup_{j=1}^{c(v)}(M \cap V_{v_j})$ (disjoint). When $v \in M$, then $(M \cap V_{v_j}) \cup (\{v\} \cup P_i)$ is not a cutset for every $j = 1, \dots, c(v)$ and M is written as $M = \{v\} \cup_{j=1}^{c(v)}(M \cap V_{v_j})$ (disjoint). Thus, (3.1) follows by considering whether the set M includes the vertex v or not. When $i = k-1$, $M \cup P_{k-1}$ becomes a cutset of the system if M includes the vertex v. Then, we have (3.2) by considering only the subsets which do not include the vertex v. This completes the proof.

Remark 3.1. By calculating $\eta_{v_0}^{(0)}(t)$ for the root v_0 of T, we have the number of subsets which are not cutsets of the system.

Remark 3.2. Similarly as Remark 2.1, we can add another type of boundary conditions; for $i = 0, 1, \dots, k-1$,

$$\eta_v^{(i)}(t) = (t+1)^{|V_v|} \quad \text{if} \quad i + |T_v| + 1 < k.$$

From Theorem 3.1 we obtain the number of minimal i-cutsequences of the consecutive system on a directed tree. We illustrate how to derive it by using the same directed tree as Example 2.1.

Example 3.1. We consider the directed tree with 18 vertices given in Fig. 2. Suppose that random variables $\xi_1, \dots, \xi_{18}$, which represent the lifetimes of the components, are given corresponding to the vertices of the directed tree, respectively. We assume that $\xi_1, \dots, \xi_{18}$ are independent and identically distributed and that the system fails if and only if consecutive three components along the direction fail. Following Theorem 3.1, we derive the number of subsets which are not cutsets of the system. Since the root v_1 has three children $\{v_2, v_3, v_4\}$, we have from (3.1)

$$\eta_{v_1}^{(0)}(t) = \eta_{v_2}^{(0)}(t)\eta_{v_3}^{(0)}(t)\eta_{v_4}^{(0)}(t) + t\eta_{v_2}^{(1)}(t)\eta_{v_3}^{(1)}(t)\eta_{v_4}^{(1)}(t).$$

Similarly, since the vertex v_2 has two children $\{v_5, v_6\}$, we have from Theorem 3.1 and Remark 3.2

$$\begin{aligned}\eta_{v_2}^{(0)}(t) &= \eta_{v_5}^{(0)}(t)\eta_{v_6}^{(0)}(t) + t\eta_{v_5}^{(1)}(t)\eta_{v_6}^{(1)}(t)\\ &= (t+1)(t+1)^3 + t((t+1)(\eta_{v_{11}}^{(0)}(t)\eta_{v_{12}}^{(0)}(t) + t\eta_{v_{11}}^{(2)}(t)\eta_{v_{12}}^{(2)}(t)))\\ &= (t+1)^4 + t((t+1)((t+1)^2 + t)).\end{aligned}$$

Similar argument implies

$$\begin{aligned}\eta_{v_3}^{(0)}(t) &= (t+1)^2\\ \eta_{v_4}^{(0)}(t) &= (t+1)^9 + t((t+1)^2 + t)((t+1) + t)((t+1)^3 + t)\\ \eta_{v_2}^{(1)}(t) &= (t+1)^4 + t(t+1)^2\\ \eta_{v_3}^{(1)}(t) &= (t+1) + t \cdot 1\\ \eta_{v_4}^{(1)}(t) &= (t+1)^9 + t(t+1)^6.\end{aligned}$$

Hence, we have

$$\begin{aligned}\eta_{v_1}^{(0)}(t) = &((t+1)^4 + t((t+1)((t+1)^2+t)))(t+1)^2 \\ &\times((t+1)^9 + t((t+1)^2+t)((t+1)+t)((t+1)^3+t)) \\ &+t((t+1)^4 + t((t+1)^2)((t+1)+t)((t+1)^9+t(t+1)^6).\end{aligned}$$

By expanding the generating function, we have that for $i = 0, 1, 2, \ldots, 15$ the numbers of the subsets which are not cutsets with i elements are $\{1, 18, 153, 802, 2867, 7351, 13912, 19771, 21310, 17488, 10893, 5084, 1729, 406, 59, 4\}$, respectively. Therefore, we have for $i = 0, 1, \ldots, 18$, $\omega_1 = 0$, $\omega_2 = 0$, $\omega_3 = 7/408$, $\omega_4 = 281/6120$, $\omega_5 = 199/2520$, $\omega_6 = 12091/111384$, $\omega_7 = 28547/222768$, $\omega_8 = 15667/116688$, $\omega_9 = 13927/109395$, $\omega_{10} = 8077/72930$, $\omega_{11} = 7805/87516$, $\omega_{12} = 265/3978$, $\omega_{13} = 7/153$, $\omega_{14} = 43/1530$, $\omega_{15} = 11/765$, $\omega_{16} = 1/204$, $\omega_{17} = 0$ and $\omega_{18} = 0$.

Similarly as in Section 2, we treat a complete m-ary tree of length $(n-1)$ as a special case. Since every vertex except for a leaf has just m children, we see that for v_1 and $v_2 \in V$ and for $i = 0, 1, 2, \ldots, k-1$, $\eta_{v_1}^{(i)}(t) = \eta_{v_2}^{(i)}(t)$ holds if $|T_{v_1}| = |T_{v_2}|$. Hence we denote for $i = 0, 1, \ldots, k-1$, $\eta_{j+1}^{(i)}(t) = \eta_v^{(i)}(t)$ when $|T_v| = j$. Then, from Theorem 3.1 we obtain

COROLLARY 3.1. *The generating functions of the numbers of minimal cutsequences in the consecutive-k-out-of-$(m^n-1)/(m-1)$:F system on the complete m-ary tree of length $(n-1)$ satisfy the recurrence relations;*

$$\begin{cases} \eta_n^{(0)}(t) = (\eta_{n-1}^{(0)}(t))^m + t(\eta_{n-1}^{(0)}(t))^m \\ \eta_{n-1}^{(1)}(t) = (\eta_{n-2}^{(0)}(t))^m + t(\eta_{n-2}^{(2)}(t))^m \\ \cdots \\ \eta_{n-k+1}^{(k-1)}(t) = (\eta_{n-k}^{(0)}(t))^m, \end{cases}$$

with boundary conditions $\eta_j^{(i)}(t) = (t+1)^{(m^j-1)/(m-1)}$ *for* $i + j < k$.

Remark 3.3. By setting $m = 1$ in Corollary 3.1, we can study the number of minimal i-cutsequences in the linear consecutive-k-out-of-n:F system with the generating function approach. When $m = 1$, the above system can be reduced to

$$\begin{cases} \eta_n^{(0)}(t) = \sum_{j=0}^{k-1} t^j \eta_{n-j-1}^{(0)}(t) & \text{if} \quad n \ge k \\ \eta_n^{(0)}(t) = (t+1)^n & \text{if} \quad 0 \le n < k. \end{cases}$$

Of course, we can solve this recurrence relations, since it is linear. However, an explicit form of the solution is well known to be

$$\eta_n^{(0)}(t) = \sum_{j=0}^{[n/k]} (-1)^j \binom{n-i+1}{j} \binom{n-kj}{n-i} t^i.$$

Here, the coefficient of t^i is the number of ways of putting i identical objects into $(n-i+1)$ different cells when each of $(n-i+1)$ different cells may contain no more than $(k-1)$ objects (cf. e.g. Riordan (1958), p. 104).

4. Computational aspects

The results obtained in this paper are intended for computation as well as for theoretical development. Theorems 2.1 and 3.1 and Corollaries 2.1 and 3.1 indeed provide algorithms for the corresponding computations. In fact, it is easy to convert the theorems to recursive procedures by using some computer algebra systems. Though we used in Sections 2 and 3 the directed tree of very small length with only 18 vertices to illustrate how to derive the generating functions, the size of the directed tree is not a problem if we input the directed tree to the computer by using standard methods like adjacency list representation or adjacency matrix representation. Indeed, we can treat the corresponding computational results to Examples 2.1 and 3.1 for the directed tree with 41 vertices in Fig. 1. The *pgf* of the distribution of "1"-runs of length 3 on the directed tree becomes a polynomial in t of degree 18, which is very long and may not be suitable for printing all the polynomial. Nevertheless, it is very easy to obtain the probabilities from the *pgf* by means of some computer algebra systems, since most computer algebra systems are excellent in expanding polynomials and in taking out the coefficients. More generally, it is also well known that recurrence relations of probabilities or moments are derived by standard methods from the *pgf* if the *pgf* is a rational function (cf. e.g. Stanley (1986), Chapter 4).

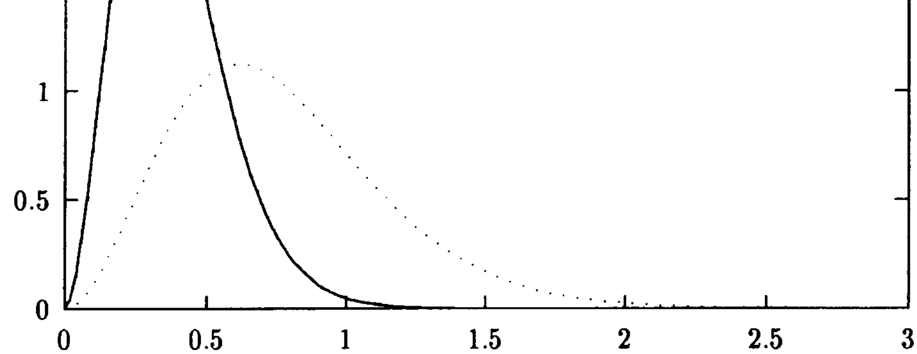

Fig. 3. Density function of the lifetime of the consecutive-3-out-of-41:F system on the directed tree.

We can also derive the lifetime of the consecutive-3-out-of-41:F system on the directed tree. By calculating the generating function of the numbers of the subsets which are not cutsets with i elements, we obtain the mixing weights $\omega_1, \dots, \omega_{41}$ exactly. In particular, if we further assume that $G(t) = 1 - \exp(-\mu t)$ (the exponential distribution with parameter μ), the density of the lifetime of the system can be written exactly. The mean and variance of the lifetime of the system are given as

$$\text{mean} = \frac{513842337225463}{1335732864265800\mu}$$

$$\text{variance} = \frac{8463579083589886828808912875 9}{2378909712906290782097399520000\mu^2}.$$

Figure 3 shows the graphs of the density functions of the lifetime of the consecutive-3-out-of-41:F system relative to the directed tree with exponential components ($\mu = 1$ and $1/2$).

References

Aki, S. (1985). Discrete distributions of order k on a binary sequence, *Ann. Inst. Statist. Math.*, **37**, 205–224.

Aki, S. (1992). Waiting time problems for a sequence of discrete random variables, *Ann. Inst. Statist. Math.*, **44**, 363–378.

Aki, S., Balakrishnan, N. and Mohanty, S. G. (1996). Sooner and later waiting time problems and failure runs in higher order Markov dependent trials, *Ann. Inst. Statist. Math.*, **48**, 773–787.

Aki, S. and Hirano, K. (1988). Some characteristics of the binomial distribution of order k and related distributions, *Statistical Theory and Data Analysis II, Proceedings of the 2nd Pacific Area Statistical Conference* (ed. K. Matusita), 211–222, North-Holland, Amsterdam.

Aki, S. and Hirano, K. (1993). Discrete distributions related to succession events in a two-state Markov chain, *Statistical Science & Data Analysis* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467–474, VSP Publishers, Amsterdam.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **47**, 225–235.

Aki, S. and Hirano, K. (1996). Lifetime distribution and estimation problems of consecutive-k-out-of-n:F systems, *Ann. Inst. Statist. Math.*, **48**, 185–199.

Balakrishnan, N., Mohanty, S. G. and Aki, S. (1997). Start-up demonstration tests under Markov dependence model with corrective actions, *Ann. Inst. Statist. Math.*, **49**, 155–169.

Balasubramanian, K., Viveros, R. and Balakrishnan, N. (1993). Sooner and later waiting time problem for Markovian Bernoulli trials, *Statist. Probab. Lett.*, **18**, 153–161.

Chao, M. T., Fu, J. C. and Koutras, M. V. (1995). Survey of reliability studies of consecutive-k-out-of-n:F & related systems, *IEEE Transactions on Reliability*, **40**, 120–127.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner and later problems for Bernoulli trials: frequency and run quotas, *Statist. Probab. Lett.*, **9**, 5–11.

Fu, J. C. (1996). Distribution theory of runs and patterns associated with a sequence of multi-state trials, *Statistica Sinica*, **6**, 957–974.

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

Godbole, A. P. and Papastavridis, S. G. (1994). *Runs and Patterns in Probability: Selected Papers*, Kluwer, Dordrecht.

Hirano, K. (1986). Some properties of the distributions of order k, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou, G. E. Bergum and A. F. Horadam), 43–53, Reidel, Dordrecht.

Hirano, K. (1994). Consecutive-k-out-of-n:F Systems, *Proc. Inst. Statist. Math.*, **42**, 45–61 (in Japanese).

Hirano, K. and Aki, S. (1993). On number of occurrences of success runs of specified length in a two-state Markov chain, *Statistica Sinica*, **3**, 313–320.

Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete distributions*, Wiley, New York.

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: a unified Markov chain approach, *Ann. Inst. Statist. Math.*, **47**, 743–766.

Lauritzen, S. L. (1996). *Graphical Models*, Clarendon Press, Oxford.

Lou, W. Y. W. (1996). On runs and longest run tests: a method of finite Markov chain imbedding, *J. Amer. Statist. Assoc.*, **91**, 1595–1601.

Mohanty, S. G. (1994). Success runs of length k in Markov dependent trials, *Ann. Inst. Statist. Math.*, **46**, 777–796.

Philippou, A. N. (1986). Distributions and Fibonacci polynomials of order k, longest runs, and reliability of consecutive-k-out-of-n:F system, *Fibonacci Numbers and Their Applications* (eds. A. N. Philippou, G. E. Bergum and A. F. Horadam), 203–227, Reidel, Dordrecht.

Philippou, A. N., Georghiou, C. and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties, *Statist. Probab. Lett.*, **1**, 171–175.

Philippou, A. N. and Makri, F. S. (1986). Successes, runs, and longest runs, *Statist. Probab. Lett.*, **4**, 101–105.

Philippou, A. N. and Muwafi, A. A. (1982). Waiting for the k-th consecutive success and the Fibonacci sequence of order k, *The Fibonacci Quart.*, **20**, 28–32.

Rajarshi, M. B. (1974). Success runs in a two-state Markov chain, *J. Appl. Probab.*, **11**, 190–192.

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

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*, Cambridge University Press, Cambridge.

Schwager, S. J. (1983). Run probabilities in sequences of Markov-dependent trials, *J. Amer. Statist. Assoc.*, **78**, 168–175.

Stanley, R. T. (1986). *Enumerative Combinatorics*, Wadsworth Publishers, Kentucky.

Uchida, M. and Aki, S. (1995). Sooner and later waiting time problems in two-state Markov chain, *Ann. Inst. Statist. Math.*, **47**, 415–433.