Methodol Comput Appl Probab (2013) 15:453–465
DOI 10.1007/s11009-011-9268-1

On Finite Markov Chain Imbedding and Its Applications

Tung-Lung Wu

Received: 3 February 2011 / Revised: 25 October 2011 /
Accepted: 16 November 2011 / Published online: 4 December 2011

Abstract The finite Markov Chain Imbedding technique has been successfully applied in various fields for finding the exact or approximate distributions of runs and patterns under independent and identically distributed or Markov dependent trials. In this paper, we derive a new recursive equation for distribution of scan statistic using the finite Markov chain imbedding technique. We also address the problem of obtaining transition probabilities of the imbedded Markov chain by introducing a notion termed Double Finite Markov Chain Imbedding where transition probabilities are obtained by using the finite Markov chain imbedding technique again. Applications for random permutation model in chemistry and coupon collector's problem are given to illustrate our idea.

Keywords Finite Markov chain imbedding · Transition matrix · Random permutation · Scan statistic

AMS 2000 Subject Classification 60E05 · 60J10

1 Introduction

Fu and Koutras (1994) proposed the finite Markov chain imbedding (FMCI) technique to derive the exact distributions associated with several runs statistics in either independent and identically distributed (i.i.d.) or Markov dependent trials. It has been shown in the literature that the FMCI technique has been successfully applied in many areas, such as reliability (Cui et al. 2010), quality control (Chang and Wu 2011), boundary crossing problem (Fu and Wu 2010). In particular, the FMCI technique has been shown to be useful in some classic random permutation problems, for example number of successions (Fu 1995) and Eulerian and Simon Newcomb numbers (Fu

T.-L. Wu (✉)
Department of Statistics, University of Manitoba, Winnipeg, MB, Canada R3T 2N2
e-mail: umwu48@cc.umanitoba.ca

et al. 1999); however, sometimes it can be very tedious using combinatorial method. Hence, the FMCI technique is often served as an alternative to the combinatorial method for finding the exact distributions.

The fundamental theory of FMCI relies on the Chapman–Kolmogorov equation, whose central idea depends on being able to imbed or turn a statistic or a random variable into a finite Markov chain, and the distribution of the statistic is obtained via transition matrices of the imbedded Markov chain. In many problems, the imbedding procedures of turning non-Markov random variables into Markov chains are implemented case by case. Fu (1996) developed a procedure called "*forward and backward principle*" to systematically carry out the imbedding procedure and obtained the exact distributions for general runs and patterns in multistate trails. The forward and backward principle demonstrates how to construct an imbedded Markov chain carrying all the information included in the original random variable or statistic.

In the literature, many authors have shown how to imbed a finite Markov chain; however, the difficulty of obtaining the transition probabilities of the imbedded Markov chain has not been addressed and studied. Without explicit transition probabilities, the FMCI technique may fail to obtain the desired distributions. To extend the flexibility and usefulness of the FMCI technique, we introduce an extension of the FMCI technique called double finite Markov chain imbedding (DFMCI). The motivation of DFMCI is the desire to resolve the difficulty mentioned above and the idea is to use the FMCI technique repeatedly to explicitly obtain the transition probabilities of the imbedded Markov chain.

This paper is organized in the following way. Section 2 provides the formulation of the FMCI technique. A new recursive equation for distribution of scan statistic using the FMCI technique is obtained, demonstrating the ability of the FMCI technique. In Section 3, we introduce the idea of DFMCI and show how to implement the procedure for two practical problems: random permutation model in chemistry and coupon collector's problem.

2 Finite Markov Chain Imbedding Technique

Definition 1 (FMCI, Fu and Lou 2003) A non-negative integer-valued random variable X_n, taking values in $\Gamma_{X_n} = \{0, 1, \ldots, \ell_n\}$, is finite Markov chain imbeddable if there exists a finite Markov chain $\{Y_t\}$ defined on a finite state space $\Omega = \{a_1, a_2, \ldots, a_m\}$ of size m with initial probability vector ξ_0 such that

$$P(X_n = x) = P(Y_n \in C_x | \xi_0) = \xi_0 \mathbf{M}^n \mathbf{U}^{'}(C_x), \quad x \in \Gamma_{X_n}, \tag{1}$$

where C_x is a subset of Ω corresponding to x, $\mathbf{U}(C_x) = \sum_{r:a_r \in C_x} \mathbf{e}_r$, $\mathbf{e}_r$ is a $1 \times m$ unit row vector corresponding the state a_r, and $\mathbf{M}$ is the transition probability matrix of the imbedded Markov chain. Note that if the Markov chain is nonhomogeneous, then the right hand side of Eq. 1 becomes the product of the transition probability matrices.

First-entry probability or absorption probability appears in many statistical problems and it is an important feature of the FMCI technique. If there is an absorbing state α in Ω, then the transition matrix $\mathbf{M}$ can be partitioned as

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{N} & \mathbf{C} \\ \hline \mathbf{0} & 1 \end{array}\right]_{m\times m}. \tag{2}$$

Define the stopping time

$$T = \inf\{t : Y_t = \alpha\},$$

then it follows from Eq. 1 that

$$P(T \leq t) = P(Y_t = \alpha) = \boldsymbol{\xi}_0 \mathbf{M}^t \mathbf{e}'_m = 1 - \boldsymbol{\xi}_0 \mathbf{N}^t \mathbf{1}^{'}, \tag{3}$$

and the mean $E(T)$ and the probability generating function $\varphi_T(s)$ of T are, respectively, given by

$$E(T) = \boldsymbol{\xi}_0(\mathbf{I} - \mathbf{N})^{-1}\mathbf{1}^{'}, \tag{4}$$

$$\varphi_T(s) = 1 + (s-1)\boldsymbol{\xi}_0(\mathbf{I} - s\mathbf{N})^{-1}\mathbf{1}^{'}, \tag{5}$$

where $\mathbf{1}^{'}$ is a column vector with all elements 1.

2.1 A New Recursive Equation for Scan Statistic

Let $\{X_i\}$ be a sequence of i.i.d. two-state $\{0, 1\}$-valued trials with probabilities p_0 and p_1, respectively. The scan statistic $S_n(w)$ of window size w is defined by

$$S_n(w) = \max_{1\leq t\leq n-w+1} S(w, t),$$

where $S(w, t) = \sum_{i=t}^{t+w-1} X_i$.

Fu (2001) studied the exact distribution of scan statistic where the probability $P(S_n(w) < s)$ is expressed as the tail probability of waiting time variable of a compound pattern which has been shown to be finite Markov chain imbeddable and studied extensively. Here, we propose another imbedding procedure which leads to a recursive equation for distribution of scan statistic. In order to construct a Markov chain $\{Y_t\}$ carrying all the information for a scan statistic of window size w, we choose to keep track of locations of successes (1's) counting backward in a window of size w at each time t. Given window size w, we define a state space

$$\Omega = \{0\} \cup \{x_1 \cdots x_j : x_j = 1, 2, \ldots, w,\ x_1 < \cdots < x_j \text{ and } j = 1, 2, \ldots, w\}.$$

We further define a finite Markov chain $\{Y_t\}$ on the state space Ω. It is clear that Y_t includes the locations of successes and total number of successes in the window of size w at time t. For example, given $n = 10$ and $w = 4$ with outcomes 0100001101, it follows the state space $\Omega = \{0, 1, 2, 12, 3, 13, 23, 123, 4, 14, 24, 124, 34, 134, 234, 1234\}$ and the realization of the Markov chain $\{Y_t : t = 0, \ldots, 10\}$ is $\{Y_0 = 0, Y_1 = 0, Y_2 = 1, Y_3 = 2, Y_4 = 3, Y_5 = 4, Y_6 = 0, Y_7 = 1, Y_8 = 12, Y_9 = 23, Y_{10} = 134\}$. We say that state 134 is of length 3 denoted by $\ell(134) = 3$ and other states apply by analogy except for state 0 which is defined to be of length zero.

In order to derive the recursive formula, we consider each state being a base-2 number and we attach each state a label obtained by transforming the base-2 number into a base-10 number plus one. For example, state 13 is attached a label $2^{1-1} + 2^{3-1} + 1 = 6$ and state 24 is assigned with a label $2^{2-1} + 2^{4-1} + 1 = 11$. In particular, state 0 receives a label 1; i.e.

state:	0	1	2	12	3	13	23	123	4	14	24	124	34	134	234	1234
label:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

By such labeling, the transition probabilities of the finite Markov chain $\{Y_t\}$ are given by

$$P(Y_t = v|Y_{t-1} = u) = \begin{cases} p_0 & \text{if } v_L = 2u_L - 1 \text{ and } u \text{ does not contain the character } w, \\ p_1 & \text{if } v_L = 2u_L \text{ and } u \text{ does not contain the character } w, \\ p_0 & \text{if } v_L = 2(u_L - 2^{w-1}) - 1 \text{ and } u \text{ contains the character } w, \\ p_1 & \text{if } v_L = 2(u_L - 2^{w-1}) \text{ and } u \text{ contains the character } w, \\ 0 & \text{otherwise}, \end{cases} \tag{6}$$

where u_L and v_L are labels associated with states u and v, respectively. It follows from Eq. 6 that the transition matrix of $\{Y_t\}$ has the form as follows: for example $w = 4$,

$$\mathbf{M} = \begin{array}{c} 0 \\ 1 \\ 2 \\ 12 \\ 3 \\ 13 \\ 23 \\ 123 \\ 4 \\ 14 \\ 24 \\ 124 \\ 34 \\ 134 \\ 234 \\ 1234 \end{array} \begin{pmatrix} p_0 & p_1 & & & & & & & & & 0 & & & & & \\ & & p_0 & p_1 & & & & & & & & & & & & \\ & & & & p_0 & p_1 & & & & & & & & & & \\ & & & & & & p_0 & p_1 & & & & & & & & \\ & & & & & & & & p_0 & p_1 & & & & & & \\ & & & & & & & & & & p_0 & p_1 & & & & \\ & & & & & & & & & & & & p_0 & p_1 & & \\ & & & & & & & 0 & & & & & & & p_0 & p_1 \\ p_0 & p_1 & & & & & & & & & & & & & & \\ & & p_0 & p_1 & & & & & & & & & & & & \\ & & & & p_0 & p_1 & & & & & & & & & & \\ & & & & & & p_0 & p_1 & & & & & & & & \\ & & & & & & & & p_0 & p_1 & & & & & & \\ & & & & & & & & & & p_0 & p_1 & & & & \\ & & & & 0 & & & & & & & & p_0 & p_1 & & \\ & & & & & & & & & & & & & & p_0 & p_1 \end{pmatrix}. \tag{7}$$

Let $a_t(v_L) = P(Y_t = v)$, it follows from Eqs. 6 and 7 that we have the following recursive equations for $a_t(v_L)$: $t = 1, \ldots, n$, and $v_L = 1, \ldots, 2^w$,

if $v_L =$ odd,

$$a_t(v_L) = p_0 a_{t-1}\left(\frac{v_L + 1}{2}\right) + p_0 a_{t-1}\left(2^{w-1} + \frac{v_L + 1}{2}\right), \tag{8}$$

if $v_L =$ even,

$$a_t(v_L) = p_1 a_{t-1}\left(\frac{v_L}{2}\right) + p_1 a_{t-1}\left(2^{w-1} + \frac{v_L}{2}\right), \tag{9}$$

with initial conditions $a_0(1) = 1$ and $a_0(v_L) = 0$ if $v_L \neq 1$.

For given $s \leq w$, the probability $P(S_n(w) < s)$ can be obtained from the above recursive equations with some modification. The event $\{S_n(w) < s\}$ represents that the Markov chain $\{Y_t\}$ never attains states of length greater than or equal to s. Thus, from Eqs. 8 and 9, it yields the following theorem.

Theorem 1 *Given n, window size w and s, we have*

$$P(S_n(w) < s) = \sum_{v_L \in A_s} a_n(v_L),$$

where A_s is a set of labels associated with states of length less than s, and $a_n(v_L)$ satisfies the following recursive equations: $t = 1, \ldots, n$, and $v_L = 1, \ldots, 2^w$,

if $v_L = odd$,

$$a_t(v_L) = p_0 a_{t-1}\left(\frac{v_L + 1}{2}\right) + p_0 a_{t-1}\left(2^{w-1} + \frac{v_L + 1}{2}\right), \qquad (10)$$

if $v_L = even$,

$$a_t(v_L) = p_1 a_{t-1}\left(\frac{v_L}{2}\right) + p_1 a_{t-1}\left(2^{w-1} + \frac{v_L}{2}\right), \qquad (11)$$

with initial conditions $a_0(1) = 1$ and $a_0(v_L) = 0$ if $v_L \neq 1$, and $a_t(v_L) = 0$ if $v_L \notin A_s$ for all $0 < t \leq n$.

From the above recursive equations, we can derive the recursive equation for $P(S_n(w) < s)$. A state of length $s-1$ at time n may be entered from a state of length s with w in it at time $n-1$, combined with $X_n = 0$. It follows that the later in our recursive equations has probability 0 and should be excluded from the probability $P(S_n(w) < s-1)$. This yields the following corollary.

Corollary 1 *For $s \geq 1$, we have*

$$P(S_n(w) < s) = P(S_{n-1}(w) < s) - p_1 \sum_{v_L \in B_s} a_{n-1}(v_L),$$

where B_s is a set of labels associated with states of length $s-1$ without w in the states.

Remark 1 Table 1 provides the probabilities $P(S_n(w) < s)$ for $n = 200$, $w = 25$ and $s = 1, 2, \ldots, 24$, under $p_0 = p_1 = 0.5$. Our result can be easily extended to Markov dependent sequences by simply replacing p_0 and p_1 by the transition probabilities of the underlying Markov chains including nonhomogeneous Markov chains. The probability generating function for $P(S_n(w) < s)$ is not provided since it can be derived straightforward from the definition of probability generating function and our recursive equations.

Table 1 Probabilities $P(S_n(w) < s)$ for $n = 200$, $w = 25$ and $s = 1, 2, \ldots, 24$, under $p_0 = p_1 = 0.5$

s	$P(S_{200}(25) < s)$	s	$P(S_{200}(25) < s)$	s	$P(S_{200}(25) < s)$	s	$P(S_{200}(25) < s)$
1	6.2230×10^{-61}	7	4.1149×10^{-22}	13	4.5828×10^{-5}	19	7.7126×10^{-1}
2	4.0210×10^{-52}	8	4.3051×10^{-18}	14	1.1503×10^{-3}	20	9.1690×10^{-1}
3	1.7588×10^{-44}	9	1.3730×10^{-14}	15	1.3381×10^{-2}	21	9.7722×10^{-1}
4	8.6885×10^{-38}	10	1.4531×10^{-11}	16	7.8447×10^{-2}	22	9.9536×10^{-1}
5	6.8185×10^{-32}	11	5.4854×10^{-9}	17	2.5636×10^{-1}	23	9.9933×10^{-1}
6	1.0801×10^{-26}	12	7.8800×10^{-7}	18	5.2588×10^{-1}	24	9.9994×10^{-1}

3 Double Finite Markov Chain Imbedding

In the previous section, we introduce some basic results for the FMCI technique. The transition matrix of the imbedded Markov chain plays an important role of this technique, as in the scan statistic problem, it is completely characterized by the transition probabilities of the imbedded Markov chain $\{Y_t\}$, p_0 and p_1. However, there are problems suffering from the difficulty of finding the transition probabilities, or finding the transition probabilities is not as easy as it is in the above scan statistic problem. We introduce a notion of DFMCI in the following. Each row of transition matrix of an imbedded Markov chain sums to one and can be considered as a distribution of some random variable. If the random variable possesses Markov property or is finite Markov chain imbeddable, then the distribution of the random variable or transition probabilities of the imbedded Markov chain can be obtained by using the FMCI technique again. In this section, we will show the details to implement FMCI and DFMCI procedures for a random permutation model in chemistry and three variations of coupon collector's problem.

3.1 Random Permutation in Chemistry

A classic hat-check problem is well-known and various variations have been proposed and studied in the literature, for example see Ross and Peköz (2007) and Scoville (1966). Brown et al. (2008) studied a variation: a model for chemical bonding process arising in chemistry. The problem can be formulated as follows: n elements are numbered from 1 to n, and randomly permuted. If the element $k+1$ is located to the right of the element k, then elements k and $k+1$ are bonded and form a cluster. Similarly, if a cluster ending with number k is followed by a cluster starting with number $k+1$, then they form a new cluster. This process is repeated until there is only one cluster remaining. A random variable T_n is defined as the number of random permutations such that only one cluster is left. We study the distribution of T_n by using DFMCI in the following.

First we define Y_m to be the number of clusters after m permutations. Clearly, $\{Y_m\}$ is a finite Markov chain defined on a state space $\Omega = \{1, 2, \ldots, n\}$. In the beginning, each number is considered as a cluster, hence state n is the initial state, and state 1 is regarded as an absorbing state α. Utilizing the FMCI technique, we have

$$P(T_n > t) = P(Y_1 \neq \alpha, Y_2 \neq \alpha, \ldots, Y_t \neq \alpha) = \boldsymbol{\xi}_0 \mathbf{N}^t \mathbf{1}', \tag{12}$$

where $\boldsymbol{\xi}_0 = (0, 0, \ldots, 0, 1)$, $\mathbf{N}$ is the transition sub-matrix in Eq. 2. Intuitively, the transition probabilities are not easy to obtain and complicated, but we will show the transition probabilities are finite Markov chain imbeddable and can be computed using the FMCI technique again. Define a random variable C_i to be the number of clusters formed after a new permutation starting from state i and it follows that with $k = i - j \geq 0$,

$$P(Y_t = j | Y_{t-1} = i) = P(C_i = k), \quad t \leq n. \tag{13}$$

It is not difficult to discover that the random variable C_i is also the number of successions of size 2 in a random permutation problem. Let $\mathscr{P}_n = \{\pi : \pi = (\pi_1, \ldots, \pi_n), \pi_i \in \{1, \ldots, n\}, i = 1, \ldots, n\}$ be the set of all permutations generated by integers $1, \ldots, n$. Any pair (π_i, π_{i+1}) is said to be a succession of size 2 in a random permutation $\boldsymbol{\pi}$ if $\pi_{i+1} = \pi_i + 1$. The number of successions of size 2 is defined by

$$X_n(\pi) = \sum_{i=1}^{n-1} I_n(i, \pi),$$

where

$$I_n(i, \pi) = \begin{cases} 1 & \text{if } \pi_{i+1} = \pi_i + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

It has been shown in Fu (1995) that by insertion procedure $X_n(\pi)$ is finite Markov chain imbeddable, and there exists a finite Markov chain $\{Y_t^C\}$ defined on a state space $\Omega^C = \{0, 1, \ldots, n-1\}$ having transition probabilities as follows:

$$P(Y_t^C = j | Y_{t-1}^C = i) = \begin{cases} \dfrac{i}{t} & \text{if } j = i - 1, \\[2ex] \dfrac{t-i-1}{t} & \text{if } j = i, \\[2ex] \dfrac{1}{t} & \text{if } j = i + 1. \end{cases} \tag{15}$$

It can be seen from Eq. 15 that $\{Y_t^C\}$ is a nonhomogeneous Markov chain having transition matrices of the form

$$\mathbf{M}_t^C = \begin{array}{c} 0 \\[1ex] 1 \\[1ex] 2 \\[1ex] \vdots \\[1ex] t-3 \\[1ex] t-2 \\[1ex] t-1 \\ \vdots \\ n-1 \end{array} \left[\begin{array}{cccccc|c} \frac{t-1}{t} & \frac{1}{t} & & \mathbf{0} & & & \\ \frac{1}{t} & \frac{t-2}{t} & \frac{1}{t} & & & & \\ & \frac{2}{t} & \frac{t-3}{t} & \frac{1}{t} & & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & & \\ & \mathbf{0} & & \frac{t-3}{t} & \frac{2}{t} & \frac{1}{t} & \\ & & & & \frac{t-2}{t} & \frac{1}{t} & \frac{1}{t} \\ \hline & & & & & & \\ & & \mathbf{0} & & & & I_{n-t+1} \\ & & & & & & \end{array} \right], \tag{16}$$

where I_{n-t+1} is the $(n-t+1) \times (n-t+1)$ identity matrix. It follows from the definition of FMCI that

$$P(C_i = k) = \boldsymbol{\xi}_0^C \left(\prod_{t=1}^{i} \mathbf{M}_t^C \right) \mathbf{e}_k^{'}, \quad k = 0, \ldots, n-1, \tag{17}$$

where $\boldsymbol{\xi}_0^C = (1, 0, \ldots, 0)$. Therefore, by using the FMCI technique again we have transition probabilities given in Eq. 17 for the transition probability sub-matrix $\mathbf{N}$ in

Table 2 The exact values and lower and upper bounds for $E(T_n)$ with $n = 100$, 500 and 1000

n	Exact	Lower bound	Upper bound
100	103.8292	102.6175	104.1947
500	505.4386	504.0866	505.8081
1000	1006.1318	1004.7232	1006.5018

Eq. 12 to calculate the distribution of T_n. Since $\{Y_m\}$ is a homogeneous Markov chain with transition matrix of the form in Eq. 2, the mean and the probability generating function can be obtained by Eqs. 4 and 5, respectively.

Remark 2 Using DFMCI, the problem of chemical bonding process can be easily extended to that if k consecutive integers are placed in an increasing order then they form a cluster. The random variable C_i is then defined as the number of successions of size k or the number of increasing k-sequences formed after a permutation starting from state i. By the same token, we can obtain the exact distribution of number of permutations needed until only one cluster is left, while it is not a simple task using combinatorial method. The detail of the imbedding procedure for the number of successions of size k can be found in Johnson and Fu (2000). Table 2 provides exact values of $E(T_n)$ and lower and upper bounds for $E(T_n)$ given by Brown et al. (2008), for $n = 100,\ 500$ and 1000.

3.2 Coupon Collector's Problem

Generalized Coupon Collector's Problem In probability theory, the classic coupon collector's problem states as follows: suppose that there are n different coupons and the collector receives one coupon each day. Every coupon is received with probability $1/n$. What is the minimum number T_n of days required in order to have all coupons collected at least once?

A variation studied by Johnson and Sellke (2010) generalizes the problem in such a way that the collector will receive $\min(K_t, n)$ number of coupons at t-th day where K_t is a random variable, taking values in $\{1, 2, \ldots\}$, with $p_i = P(\min(K_t, n) = i)$, the probability of receiving i coupons at t-th day. Define X_t to be the number of different coupons collected after t-th day. Clearly, $\{X_t\}$ is a Markov chain defined on a state space $\Omega = \{0, 1, \ldots, n\}$, where n is an absorbing state. The transition probabilities of $\{X_t\}$ are in fact weighted averages of hypergeometric distributions given as follows:

$$P(X_{t+1} = k | X_t = j) = \sum_{i=1}^{n} p_i \frac{\binom{n-j}{k-j}\binom{j}{i-(k-j)}}{\binom{n}{i}}. \tag{18}$$

We take this example as an illustration for our method to solve the problem without knowing the transition probabilities given in Eq. 18. We define a random variable D_j as the number of different coupons collected in a day given j coupons being already collected. It turns out that the distribution of D_j provides the transition probabilities of $\{X_t\}$ starting at state j. We then show that D_j is conditional finite Markov chain imbeddable. Given that i coupons will be received on (t+1)-st day, we define $Z_m(i)$ as the number of different coupons collected until m-th coupons received at (t+1)-st day for $m \le i$. Then $\{Z_m(i)\}$ is a finite Markov chain defined on a state space $\Omega = \{0, 1, \ldots, i\}$ having transition probabilities as follows: for $m - x \leq j$ and $x \leq n - j$,

$$P(Z_{m+1}(i) = y | Z_m(i) = x) = \begin{cases} \dfrac{j - (m - x)}{n - m} & \text{if } y = x, \\ \\ \dfrac{(n - j) - x}{n - m} & \text{if } y = x + 1. \end{cases} \tag{19}$$

Therefore, we have

$$P(D_j = d | K_{t+1} = i) = \boldsymbol{\xi}_0 \mathbf{M}_i^i(z) \mathbf{e}_d^{'}, \tag{20}$$

and

$$P(D_j = d) = \sum_{i=1}^{n} p_i \boldsymbol{\xi}_0 \mathbf{M}_i^i(z) \mathbf{e}_d^{'}, \tag{21}$$

where $\boldsymbol{\xi}_0 = (1, 0, \ldots, 0)$, $\mathbf{M}_i(z)$ is the transition matrix of $\{Z_m(i)\}$ with transition probabilities obtained from Eq. 19. Hence, it follows that

$$P(X_{t+1} = k | X_t = j) = \sum_{i=1}^{n} p_i \boldsymbol{\xi}_0 \mathbf{M}_i^i(z) \mathbf{e}_{k-j}^{'}, \tag{22}$$

and

$$P(T_n \leq t) = \boldsymbol{\xi}_0 \mathbf{M}^t \mathbf{e}_n^{'} \quad \text{or} \quad P(T_n > t) = \boldsymbol{\xi}_0 \mathbf{N}^t \mathbf{1}^{'},$$

where $\mathbf{e}_n$ is an unit row vector corresponding to state n and $\mathbf{M}$ is the transition probability matrix of $\{X_t\}$, whose transition probabilities are given in Eq. 22. By the same token, we can obtain the mean and probability generating function of T_n from Eqs. 4 and 5.

Weighted Coupon Collector's Problem Next we consider a variation—weighted coupon collector's problem. Suppose there are n coupons, the collector receives one coupon a day and the chances of receiving each coupon are not equal, instead with probability p_i to receive coupon i. The ability of the FMCI technique is not only able to derive the exact distribution through transition probability matrices, but also able to derive other useful identities such as recursive equations. In this application, we derive a recursive equation for the distribution of T_n through the structure of transition probability matrix of an imbedded Markov chain. Since the chances of receiving each coupon are different, we have to record the coupons we have already collected to know the probability of next transition. We define a finite Markov chain $\{Y_t\}$ on a state space

$$\Omega = \{0, 1, \ldots, n, 12, 13, \ldots, (n-1)n, \ldots, 123 \cdots n\},$$

where state $123 \cdots n$ is an absorbing state. Then $Y_t = 1346$ means that coupons 1, 3, 4 and 6 are collected until t-th day and we say the state is of length 4 denoted by $\ell(1346) = 4$. There are total 2^n states in the state space. At next day, The number of collected coupons can be the same or increase by one, and then it follows from the definition that the transition probabilities from state $u = i_1 \cdots i_{j_{t-1}}$ at time $t-1$ to states $v = i_1 \cdots i_{j_{t-1}}$ or $i_1 \cdots i_{j_{t-1}} i_{j_t}$ at time t are given by

$$P(Y_t = v | Y_{t-1} = u) = \begin{cases} p_{i_{j_t}} & \text{if } \ell(v) = \ell(u) + 1, \\ \\ \sum_{s=1}^{j_{t-1}} p_{i_s} & \text{if } \ell(v) = \ell(u). \end{cases} \tag{23}$$

The transition matrix $\mathbf{M}$ is of the form

$$\mathbf{M} = \begin{bmatrix} 0 & \mathbf{A}(0,1) & & & & & \\ & \mathbf{A}(1,1) & \mathbf{A}(1,2) & & & & \\ & & \mathbf{A}(2,2) & \mathbf{A}(2,3) & & & \\ & & & \ddots & \ddots & & \\ & & & & \mathbf{A}(n-1,n-1) & \mathbf{A}(n-1,n) \\ & & & & \mathbf{0} & 1 \end{bmatrix}, \tag{24}$$

where the block $\mathbf{A}(x-1, x)$ stands for the transition probabilities from states of length x-1 to states of length x. Let $\tilde{\mathbf{x}}$ denote the vector consisting of all states of length x and $\boldsymbol{\alpha}_t(x) = (P(Y_t = (\tilde{\mathbf{x}})_1), P(Y_t = (\tilde{\mathbf{x}})_2), \ldots, P(Y_t = (\tilde{\mathbf{x}})_{\binom{n}{x}}))$, where $(\tilde{\mathbf{x}})_i$ stands for the i-th component in $\tilde{\mathbf{x}}$. It is not difficult to see from backward matrix multiplication that the following recursive equations hold: $t = 1, 2, \ldots,$

$$\begin{aligned} &\boldsymbol{\alpha}_0(0) = 1; \boldsymbol{\alpha}_0(x) = 0; \boldsymbol{\alpha}_t(0) = 0, \\ &\boldsymbol{\alpha}_t(x) = \boldsymbol{\alpha}_{t-1}(x-1)\mathbf{A}(x-1, x) + \boldsymbol{\alpha}_{t-1}(x)\mathbf{A}(x, x), \quad x = 1, \ldots, n-1, \\ &\boldsymbol{\alpha}_t(n) = \boldsymbol{\alpha}_{t-1}(n-1)\mathbf{A}(n-1, n) + \boldsymbol{\alpha}_{t-1}(n). \end{aligned} \tag{25}$$

The blocks $\mathbf{A}(x-1, x)$ or $\mathbf{A}(x, x)$ can be easily obtained from Eq. 23. Of course, we can use the unified formula to compute the probability

$$P(T > t) = P(Y_t \neq n) = \boldsymbol{\xi}_0 \mathbf{N}^t \mathbf{1}^{'}.$$

Coupon Collector's Problem with Bonus Another variation of coupon collector's problem with bonus is asking the following question: suppose there are b bonus coupons and ℓ ordinal coupons and the collector will receive only one coupon a day. If the collector receives a bonus coupon, then he gets another coupon immediately. In one day, the collector can get at most all bonus coupons and one ordinal coupon. What is the minimum number T of days needed to collect all bonus and ordinal coupons.

By the same token in the previous section, we only need to make slight modification for this question. We define $Y_t = (b_t, \ell_t)$, where b_t and ℓ_t are the number of different bonus and ordinal coupons collected, respectively, after t-th day. Then $\{Y_t\}$ is a Markov chain defined on a state space $\Omega = \{(i, j) : 0 \leq i \leq b, 0 \leq j \leq \ell\}$, where (b, ℓ) is an absorbing state. The bonus coupons make the problem more complicated and transition probabilities hard to find. In a day, the collector would possibly get infinitely many coupons but only maximum $b+1$ different coupons. Given that b_t bonus coupons and ℓ_t ordinal coupons are collected until t-th day, we define $Z_i = (A_i^b, A_i^\ell)$ where A_i^b and A_i^ℓ are the number of different bonus and ordinal coupons collected, respectively, up to i-th coupons received at $(t+1)$-st day. Obviously, $\{Z_i\}$ is a Markov chain defined on a state space $\Omega_z = \{(0,0), (1,0), \ldots, (b,0), (0,0^*), (1,0^*), \ldots, (b,0^*), (0,1), (1,1), \ldots, (b,1)\}$.

Note that in the state space we have 0 and 0* in order to identify that the collector received an owned bonus coupon as 0 or an owned ordinal coupon as 0*. For example, starting from state (0,0), (i) if the collector receives an owned bonus coupon, then the next state is (0,0), and the transition is denoted by $(0, 0) \to (0, 0)$; (ii) if he receives a different bonus coupon, then the next state is (1,0); (iii) if he receives an owned ordinal coupon, then the next state is (0,0*); (iv) if he receives a different ordinal coupon, then the next state is (0,1). The transition probabilities of $\{Z_i\}$ are given as follows:

$$P(Z_{i+1} = (z^b_{i+1}, z^\ell_{i+1}) | Z_i = (z^b_i, z^\ell_i)) = \begin{cases} \dfrac{b_t + z^b_i}{b + \ell} & \text{if } (z^b_i, 0) \to (z^b_i, 0), \\ \dfrac{b - b_t - z^b_i}{b + \ell} & \text{if } (z^b_i, 0) \to (z^b_i + 1, 0), \\ \dfrac{\ell_t}{b + \ell} & \text{if } (z^b_i, 0) \to (z^b_i, 0^*), \\ \dfrac{\ell - \ell_t}{b + \ell} & \text{if } (z^b_i, 0) \to (z^b_i, 1), \\ 1 & \text{if } (z^b_i, 1) \to (z^b_i, 1), \\ 1 & \text{if } (z^b_i, 0^*) \to (z^b_i, 0^*), \\ 0 & \text{otherwise}, \end{cases} \tag{26}$$

for $z^b_i = 0, 1, \ldots, b$. It follows that the transition probabilities of $\{Y_t\}$ are given by

$$\begin{aligned} &P(Y_{t+1} = (b_{t+1}, \ell_{t+1}) | Y_t = (b_t, \ell_t)) \\ &= \begin{cases} \lim_{m\to\infty} \boldsymbol{\xi}_z \mathbf{M}^m_z (\mathbf{e}'_{(b_{t+1}-b_t,0)} + \mathbf{e}'_{(b_{t+1}-b_t,0^*)}), & \text{if } \ell_{t+1} - \ell_t = 0, \\ \lim_{m\to\infty} \boldsymbol{\xi}_z \mathbf{M}^m_z \mathbf{e}'_{(b_{t+1}-b_t,1)}, & \text{if } \ell_{t+1} - \ell_t = 1, \end{cases} \end{aligned} \tag{27}$$

where $\boldsymbol{\xi}_z = (1, 0, \ldots, 0)$, $\mathbf{M}_z$ is the transition probability matrix of $\{Z_i\}$ with transition probabilities given in Eq. 26. It follows that

$$P(T \le t) = \boldsymbol{\xi}_0 \mathbf{M}^t \mathbf{e}'_{(b,\ell)},$$

where $\boldsymbol{\xi}_0 = (1, 0, \ldots, 0)$ and $\mathbf{M}$ is the transition matrix of $\{Y_t\}$ with transition probabilities given in Eq. 27.

Remark 3 Selected expectations and cumulative probabilities are given in Table 3. Using DFMCI technique, the coupon collector's problem with bonus can be generalized to that the collector can receive a random number K_i coupons at i-th day with some modification in the imbedded Markov chain similar to the ordinal coupon collector's problem in this section.

Table 3 Expectations $E(T)$ and cumulative probabilities $P(T \leq t)$ for selected b, ℓ and t

b	ℓ	t	$P(T \leq t)$	b	ℓ	t	$P(T \leq t)$	b	ℓ	$E(T)$
5	5	5	4.9783×10^{-3}	10	15	15	5.6151×10^{-8}	5	5	15.1448
		6	2.1767×10^{-2}			16	5.9453×10^{-7}	5	10	33.5156
		7	5.4742×10^{-2}			17	3.3405×10^{-6}	5	15	54.2161
		8	1.0432×10^{-1}			18	1.3248×10^{-5}	5	20	76.5192
		9	1.6771×10^{-1}			19	4.1628×10^{-5}	10	10	36.4774
		10	2.4051×10^{-1}			20	1.1031×10^{-4}	10	20	80.2331
		11	3.1801×10^{-1}			21	2.5630×10^{-4}	10	50	234.1602
		12	3.9606×10^{-1}			22	5.3608×10^{-4}	20	10	40.6165
		13	4.7147×10^{-1}			23	1.0287×10^{-3}	20	50	241.9276
		14	5.4204×10^{-1}			24	1.8369×10^{-3}	30	50	248.6490
		15	6.0649×10^{-1}			25	3.0860×10^{-3}	50	50	259.8689

4 Summary

We have demonstrated the ability of the FMCI technique by showing that the FMCI technique can be used to derive other useful identities, for example a new recursive equation for distribution of scan statistic in Section 2.1. Our recursive equations are efficient and we can compute the probabilities for moderate large window size w and s and large n.

It does not receive much attention on how to obtain transition probabilities of the imbedded Markov chain in the literature. A notion called DFMCI is introduced to compute the transition probabilities of the imbedded Markov chain by using the FMCI technique repeatedly. Two common applications are used to illustrate our idea and it also shows the flexibility of the FMCI technique in such a way that under different settings, the imbedding procedure can be applied with some minor modifications, for example the three variations of coupon collector's problems.

Acknowledgements The author wishes to thank the editor and referee for careful reading of our manuscript and helpful comments and suggestions.

References

Brown M, Peköz EA, Ross SM (2008) A random permutation model arising in chemistry. J Appl Probab 45:1060–1070

Chang YM, Wu TL (2011) On average run lengths of control charts for autocorrelated processes. Methodol Comput Appl Probab 13:419–431

Cui L, Xu Y, Zhao X (2010) Developments and applications of the finite Markov Chain imbedding approach in reliability. IEEE Trans Reliab 59:685–690

Fu JC (1995) Exact and limiting distributions of the number of successions in a random permutation. Ann Inst Stat Math 47:435–446

Fu JC (1996) Distribution theory of runs and patterns associated with a sequence of multi-state trials. Stat Sin 6:957–974

Fu JC (2001) Distribution of the scan statistic for a sequence of bistate trials. J Appl Probab 38:908–916

Fu JC, Koutras MV (1994) Distribution theory of runs: a Markov chain approach. J Am Stat Assoc 89:1050–1058

Fu JC, Lou WYW (2003) Distribution theory of runs and patterns and its applications. World Scientific, River Edge

Fu JC, Wu TL (2010) Linear and nonlinear boundary crossing probabilities for brownian motion and related processes. J Appl Probab 47:1058–1071
Fu JC, Lou WYW, Wang YJ (1999) On the exact distributions of Eulerian and Simon Newcomb numbers associated with random permutations. Stat Probab Lett 42:115–125
Johnson BC, Fu JC (2000) The distribution of increasing l-sequences in random permutations: a Markov chain approach. Stat Probab Lett 49:337–344
Johnson BC, Sellke TM (2010) On the number of i.i.d. samples required to observe all of the balls in an urn. Methodol Comput Appl Probab 12:139–154
Ross SM, Peköz EA (2007) A second course in probability. Pekozbooks, Boston
Scoville R (1966) Mathematical notes: the Hat-Check problem. Am Math Mon 73:262–265