数理解析研究所講究録 1308 巻 2003 年 29-38

Generalized Pólya urn models and related distributions

統計数理研究所 井上 潔司 (Kiyoshi Inoue)
The Institute of Statistical Mathematics

Abstract. In this paper, we consider a Pólya urn model containing balls of m different labels under a general replacement scheme, which is characterized by an $m \times m$ addition matrix of integers without constraints on the values of these m^2 integers other than non-negativity. This urn model includes some important urn models treated before. By a method based on the probability generating functions, we consider the exact joint distribution of the numbers of balls with particular labels which are drawn within n draws. As a special case, for $m = 2$, the univariate distribution, the probability generating function and the expected value are derived exactly.

We present methods for obtaining the probability generating functions and the expected values for all n exactly, which are very simple and suitable for computation by computer algebra systems. The results presented here develop a general workable framework for the study of Pólya urn models and attract our attention to the importance of the exact analysis. Our attempts are very useful for understanding non-classical urn models. Finally, numerical examples are also given in order to illustrate the feasibility of our results.

Key words and phrases: Pólya urn, replacement scheme, addition matrix, probability generating functions, expected value.

1 Introduction

Urn models have been among the most popular probabilistic schemes and have received considerable attention in the literature (see Johnson, Kotz and Balakrishnan (1997), Feller (1968)). The Pólya urn was originally applied to problems dealing with the spread of a contagious disease (see Johnson and Kotz (1977), Marshall and Olkin (1993)).

We describe the Pólya urn scheme briefly. From an urn containing α_1 balls labeled 1 and α_2 balls labeled 2, a ball is drawn, its label is noted and the ball is returned to the urn along with additional balls depending on the label of the ball drawn; If a ball labeled i $(i = 1, 2)$ is drawn, a_{ij} balls labeled j $(j = 1, 2)$ are added. This scheme is characterized by the following 2×2 addition matrix of integers, $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$; whose rows are indexed by the label selected and whose columns are indexed by the label of the ball added.

Several Pólya urn models have been studied by many authors in the various addition matrices, which generate many fruitful results. The case of the classical Pólya urn model $(a_{11} = a_{22},\ a_{12} = a_{21} = 0)$ was studied earlier and a detailed discussion can be found in Johnson and Kotz (1977). In the case of $a_{11} = a_{22}$, $a_{12} = a_{21} = 0$, Aki and Hirano (1988) obtained the Pólya distribution of order k. In the case of $a_{ii} = c$, $a_{ij} = 0$ for $i \neq j$ $(i, j = 0, 1, ..., m)$, Inoue and Aki (2000) considered the waiting time problem for the first occurrence of a pattern in the sequence obtained by an $(m+1) \times (m+1)$ Pólya urn scheme.

In the case of $a_{11} = a_{22}$, $a_{12} = a_{21}$, Friedman (1949) obtained the moment generating function of the total number of balls with a particular label remaining in the urn after n draws; Friedman's urn can be used to model the growth of leaves in recursive trees (see also Mahmoud and Smythe (1991)). In the case of $a_{11} + a_{12} = a_{21} + a_{22}$, Bagchi and Pal (1985) showed an interesting example of Pólya urn scheme applied to data structures in computer. (Gouet (1989,1993) corrected some of the statements made by Bagchi and Pal (1985)). In a $p \times p$ Pólya urn scheme (constant row sums allowing negative entries on the diagonal, but having several constraints on the eigenvalue structure), Smythe (1996) considered a central limit theorem.

One interest has been focused on the exact distribution of the total numbers of balls with particular labels remaining in the urn after n draws, or the exact distribution of the numbers of balls with particular labels which are drawn within n draws from the urn. Their derivation involves a combinatorial method of counting paths representing a realization of the urn development.

For a long time, most investigations have been made under the special structure of the constant addition matrix with constant row sums, which implies a steady linear growth of the urn size. The reason for the imposition of this constraint is mathematical convenience; Urn schemes where the constraint is imposed are generally much simpler to analyze than those where it was not imposed.

Recently, Kotz, Mahmoud and Robert (2000) attempted to treat a Pólya urn model containing 2 different labels according to a general replacement scheme, and pointed out that no constraint case is considerably more challenging even in 2×2 case. That is, the exact distribution of the number of balls with a particular label which are drawn within n draws is rather convoluted and such an exact distribution is rather unwieldy for large n for numerical computation.

Our purpose in the present paper is to develop a general workable framework for the exact distribution theory for Pólya urn models mentioned before and to emphasize the importance of the exact analysis. The approach is to solve a system of equations of conditional probability generating functions (p.g.f.'s). Then, the probability functions and moments are derived from an expansion of the solution regardless of whether or not the constraint is imposed.

In this paper, a Pólya urn model containing balls of m different labels and characterized by a general replacement scheme is considered, which include some important models treated before. We consider the exact joint distribution of the numbers of balls with particular labels which are drawn within n draws. As a special case, a univariate distribution is derived from a Pólya urn model containing balls of 2 different labels.

For the derivation of the main part of the results, we use the method based on the conditional p.g.f.'s. This method was introduced by Ebneshahrashoob and Sobel (1990), and was developed by Aki and Hirano(1993, 1999), Aki, Balakrishnan and Mohanty (1996). The procedure is very simple and suitable for computation by computer algebra systems. Furthermore, we propose a method for the Pólya urn model. It is a recurrence for obtaining the expected values for all n, which is derived from the system of equations of conditional p.g.f.'s.

The rest of this paper is organized in the following ways. In Section 2, a Pólya urn model containing balls of m different labels is introduced, which is characterized by the general replacement scheme. As a special case, a univariate distribution is derived from a Pólya urn model containing balls of 2 different labels. We give a method for the Pólya urn models. It is a recurrence for obtaining the expected values for all n. In Section 3, numerical examples are given in order to illustrate the feasibility of our main results.

2 The models

In this section, we consider a Pólya urn model characterized by an $m \times m$ addition matrix. As a special case, for $m = 2$, the univariate distribution, the probability generating function and the expected value are derived exactly.

2.1 The Pólya urn model containing m different labels

From an urn containing α_1 balls labeled 1, α_2 balls labeled 2, ..., α_m balls labeled m, a ball is chosen at random, its label is noted and the ball is returned to the urn along with additional balls according to the addition matrix of non-negative integers, $A = (a_{ij})$ $i, j = 1, ..., m$, whose rows are indexed by the label of the ball chosen and whose columns are indexed by the label of the ball added. Always starting with the newly constituted urn, this experiment is continued n times. Let $Z_1, Z_2, ..., Z_n$ be a sequence obtained by the above scheme, which take values in a finite set $B = \{1, 2, ..., m\}$. Let r be a positive integer such that $1 \leq r \leq 2^m - 1$ and let $B_1, B_2, ..., B_r$ be subsets of B, where $B_i \neq \emptyset$ and $B_i \neq B_j$ for $i \neq j$. Then, we define the numbers of balls whose labels belong to the subsets B_i $(i = 1, ..., r)$ which are drawn within n draws by $X_n^{(i)} = \sum_{j=1}^{n} I_{B_i}(Z_j)$ $(i = 1, ..., r)$, where $I_{B_i}(\cdot)$ $(i = 1, ..., r)$ means the indicator function of the subset B_i.

In the sequel, we will obtain the p.g.f. $E[t_1^{X_n^{(1)}} t_2^{X_n^{(2)}} \cdots t_r^{X_n^{(r)}}]$ of the joint distribution of $(X_n^{(1)}, X_n^{(2)}, ..., X_n^{(r)})$. Hereafter, we denote the urn composition and the total of the balls in the urn by $\boldsymbol{b} = (\alpha_1, \alpha_2, ..., \alpha_m)$ and $|\boldsymbol{b}| = \alpha_1 + \alpha_2 + \cdots + \alpha_m$, respectively. We denote the ith row of the addition matrix A by $\boldsymbol{a}_i = (a_{i1}, a_{i2}, ..., a_{im})$. Needless to say, $\alpha_i \geq 0$ $(i = 1, ..., m)$ and $|\boldsymbol{b}| \neq 0$ are assumed throughout this paper.

Suppose that we have an urn composition $\boldsymbol{b} = (\alpha_1, \alpha_2, ..., \alpha_m)$ after ℓ $(\ell = 0, 1, ..., n)$ draws. Then, we denote by $\phi_{n-\ell}(\boldsymbol{b}; \boldsymbol{t})$ the p.g.f. of the conditional distribution of the numbers of balls whose labels belong to the subsets B_i $(i = 1, ..., r)$ which are drawn within $(n - \ell)$ draws, where $\boldsymbol{t} = (t_1, ..., t_r)$.

Theorem 2.1 *From the definitions of* $\phi_{n-\ell}(\boldsymbol{b}; \boldsymbol{t})$ $(\ell = 0, 1, ..., n)$, *we have the following* *system of the equations;*

$$\text{(2.1)}\qquad \phi_n(\boldsymbol{b};\boldsymbol{t}) = \sum_{i=1}^{m} \frac{\alpha_i}{|\boldsymbol{b}|} \boldsymbol{t}^{\mathbf{I_B}(i)} \phi_{n-1}(\boldsymbol{b}+\boldsymbol{a}_i;\boldsymbol{t}),$$

$$\text{(2.2)}\qquad \phi_{n-\ell}(\boldsymbol{b};\boldsymbol{t}) = \sum_{i=1}^{m} \frac{\alpha_i}{|\boldsymbol{b}|} \boldsymbol{t}^{\mathbf{I_B}(i)} \phi_{n-\ell-1}(\boldsymbol{b}+\boldsymbol{a}_i;\boldsymbol{t}), \qquad \ell = 1,2,...,n-1,$$

$$\text{(2.3)}\qquad \phi_0(\boldsymbol{b};\boldsymbol{t}) = 1,$$

where, $\boldsymbol{t}^{\mathbf{I_B}(i)} = t_1^{I_{B_1}(i)} t_2^{I_{B_2}(i)} \cdots t_r^{I_{B_r}(i)}$.

Proof. It is easy to see that $\phi_0(\boldsymbol{b};\boldsymbol{t}) = 1$ by the definition of the p.g.f.. Suppose that the urn composition is $\boldsymbol{b} = (\alpha_1, \alpha_2, ..., \alpha_m)$ after ℓ $(\ell = 0, 1, ..., n-1)$ draws. Then, the p.g.f. of the conditional distribution of the numbers of balls whose labels belong to the subsets B_j $(j = 0, ..., r)$ which are drawn within $(n-\ell)$ draws is $\phi_{n-\ell}(\boldsymbol{b};\boldsymbol{t})$ $(\ell = 0, 1, ..., n-1)$. We should consider the condition of one-step ahead from every condition. Given the condition we observe the $(\ell+1)$-th draw. For every $i = 1, ..., m$, the probability that we draw the ball labeled i is $\alpha_i/|\boldsymbol{b}|$. If we have the ball labeled i $(i = 1, ..., m)$, then the p.g.f. of the conditional distribution of the numbers of balls whose labels belong to the subsets B_j $(j = 0, ..., r)$ which are drawn within $(n-\ell-1)$ draws is $\phi_{n-\ell-1}(\boldsymbol{b}+\boldsymbol{a}_i;\boldsymbol{t})$ $(\ell = 0, 1, ..., n-1)$. Therefore, we obtain the equations (2.1) and (2.2). □

Example 2.1 Assume that $B = \{1,2,3,4\}$, $B_1 = \{2,4\}$, $B_2 = \{3,4\}$, $\boldsymbol{t} = (t_1, t_2)$ and the addition matrix is equal to the 4×4 zero matrix. Suppose that we have an urn composition $\boldsymbol{b} = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ after ℓ $(\ell = 0, 1, ..., n)$ draws. Then, we denote by $\phi_{n-\ell}(\boldsymbol{b};\boldsymbol{t})$ the p.g.f. of the conditional distribution of the numbers of balls whose labels belong to the subsets B_1, B_2 which are drawn within $(n-\ell)$ draws. Then, we have the following system of the equations;

$$\text{(2.4)}\qquad \phi_{n-\ell}(\boldsymbol{b};t_1,t_2) = \left(\frac{\alpha_1}{|\boldsymbol{b}|} + \frac{\alpha_2}{|\boldsymbol{b}|}t_1 + \frac{\alpha_3}{|\boldsymbol{b}|}t_2 + \frac{\alpha_4}{|\boldsymbol{b}|}t_1t_2\right)\phi_{n-\ell-1}(\boldsymbol{b};t_1,t_2), \qquad \ell = 0,1,...,n-1,$$

$$\text{(2.5)}\qquad \phi_0(\boldsymbol{b};t_1,t_2) = 1.$$

Under an initial urn composition $\boldsymbol{b}_0 = (\alpha_{01}, \alpha_{02}, \alpha_{03}, \alpha_{04})$, we get

$$\text{(2.6)}\qquad \phi_n(\boldsymbol{b}_0;t_1,t_2) = \left(\frac{\alpha_{01}}{|\boldsymbol{b}_0|} + \frac{\alpha_{02}}{|\boldsymbol{b}_0|}t_1 + \frac{\alpha_{03}}{|\boldsymbol{b}_0|}t_2 + \frac{\alpha_{04}}{|\boldsymbol{b}_0|}t_1t_2\right)^n.$$

In this example, if the labels 1, 2, 3, 4 are regarded as $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, respectively, the equation (2.6) is the p.g.f. of joint distribution of the number of balls with the first label 1 and the number of balls with the second label 1 which are drawn within n draws. The distribution is called the bivariate binomial distribution (see Kocherlakota (1989), Marshall and Olkin (1985)).

2.2 The Pólya urn model containing 2 different labels

As a special case, for $m = 2$, we study the Pólya urn model containing 2 different labels. Assume that $B = \{1, 2\}$, $B_1 = \{2\}$ and $A = (a_{ij})$ $i, j = 1, 2$. Let $Y_n = \sum_{i=1}^{n} I_{B_1}(Z_i)$. Suppose that we have an urn composition $\boldsymbol{b} = (\alpha_1, \alpha_2)$ after ℓ $(\ell = 0, 1, ..., n)$ draws. Then, we denote by $\psi_{n-\ell}(\boldsymbol{b}; t_1)$ the p.g.f. of the conditional distribution of the number of balls labeled 2 which are drawn within $(n - \ell)$ draws. From Theorem 2.1, we have the following Corollary 2.1.

Corollary 2.1 *From the definitions of* $\psi_{n-\ell}(\boldsymbol{b}; t_1)$ $(\ell = 0, 1, ..., n)$*, we have the following system of the equations;*

(2.7) $\psi_n(\boldsymbol{b}; t_1) = \frac{\alpha_1}{|\boldsymbol{b}|}\psi_{n-1}(\boldsymbol{b} + \boldsymbol{a}_1; t_1) + \frac{\alpha_2}{|\boldsymbol{b}|}t_1\psi_{n-1}(\boldsymbol{b} + \boldsymbol{a}_2; t_1)$,

(2.8) $\psi_{n-\ell}(\boldsymbol{b}; t_1) = \frac{\alpha_1}{|\boldsymbol{b}|}\psi_{n-\ell-1}(\boldsymbol{b} + \boldsymbol{a}_1; t_1) + \frac{\alpha_2}{|\boldsymbol{b}|}t_1\psi_{n-\ell-1}(\boldsymbol{b} + \boldsymbol{a}_2; t_1)$, $\ell = 1, 2, ..., n - 1$,

(2.9) $\psi_0(\boldsymbol{b}; t_1) = 1$.

□

We will solve the system of the equations (2.7), (2.8) and (2.9) under an initial urn composition $\boldsymbol{b}_0 = (\alpha_{01}, \alpha_{02})$. First, we note that the above equation (2.7) can be written in matrix form as

$$\begin{aligned}
\psi_n(\boldsymbol{b}_0; t_1) &= \frac{\alpha_{01}}{\alpha_{01} + \alpha_{02}}\psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_1; t_1) + \frac{\alpha_{02}}{\alpha_{01} + \alpha_{02}}t_1\psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_2; t_1), \\
&= \left(\frac{\alpha_{01}}{\alpha_{01} + \alpha_{02}} \quad \frac{\alpha_{02}}{\alpha_{01} + \alpha_{02}}t_1\right)\begin{pmatrix} \psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_1; t_1) \\ \psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_2; t_1) \end{pmatrix}, \\
&= C_1(t_1)\boldsymbol{\psi}_{n-1}(t_1), \quad \text{(say)}.
\end{aligned}$$

Next, for $\ell = 1$, we write the equation (2.8) as

$$\begin{aligned}
\psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_1; t_1) &= \frac{\alpha_{01} + a_{11}}{\alpha_{01} + \alpha_{02} + a_{11} + a_{12}}\psi_{n-2}(\boldsymbol{b}_0 + 2\boldsymbol{a}_1; t_1) \\
&\quad + \frac{\alpha_{02} + a_{12}}{\alpha_{01} + \alpha_{02} + a_{11} + a_{12}}t_1\psi_{n-2}(\boldsymbol{b}_0 + \boldsymbol{a}_1 + \boldsymbol{a}_2; t_1), \\
\psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_2; t_1) &= \frac{\alpha_{01} + a_{21}}{\alpha_{01} + \alpha_{02} + a_{21} + a_{22}}\psi_{n-2}(\boldsymbol{b}_0 + \boldsymbol{a}_1 + \boldsymbol{a}_2; t_1) \\
&\quad + \frac{\alpha_{02} + a_{22}}{\alpha_{01} + \alpha_{02} + a_{21} + a_{22}}t_1\psi_{n-2}(\boldsymbol{b}_0 + 2\boldsymbol{a}_2; t_1),
\end{aligned}$$

or, equivalently,

$$\begin{pmatrix} \psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_1; t_1) \\ \psi_{n-1}(\boldsymbol{b}_0 + \boldsymbol{a}_2; t_1) \end{pmatrix} = \begin{pmatrix} \frac{\alpha_{01}+a_{11}}{\alpha_{01}+\alpha_{02}+a_{11}+a_{12}} & \frac{\alpha_{02}+a_{12}}{\alpha_{01}+\alpha_{02}+a_{11}+a_{12}}t_1 & 0 \\ 0 & \frac{\alpha_{01}+a_{21}}{\alpha_{01}+\alpha_{02}+a_{21}+a_{22}} & \frac{\alpha_{02}+a_{22}}{\alpha_{01}+\alpha_{02}+a_{21}+a_{22}}t_1 \end{pmatrix} \cdot \begin{pmatrix} \psi_{n-2}(\boldsymbol{b}_0 + 2\boldsymbol{a}_1; t_1) \\ \psi_{n-2}(\boldsymbol{b}_0 + \boldsymbol{a}_1 + \boldsymbol{a}_2; t_1) \\ \psi_{n-2}(\boldsymbol{b}_0 + 2\boldsymbol{a}_2; t_1) \end{pmatrix}.$$

We write $\psi_{n-1}(t_1) = C_2(t_1)\psi_{n-2}(t_1)$. For non-negative integers ℓ_1, ℓ_2 such that $\ell_1+\ell_2 = \ell$, let

$$\psi_{n-\ell}(t_1) = \begin{pmatrix} \psi_{n-\ell}(b_0 + \ell a_1; t_1) \\ \psi_{n-\ell}(b_0 + (\ell-1)a_1 + a_2; t_1) \\ \psi_{n-\ell}(b_0 + (\ell-2)a_1 + 2a_2; t_1) \\ \vdots \\ \psi_{n-\ell}(b_0 + \ell_1 a_1 + \ell_2 a_2; t_1) \\ \vdots \\ \psi_{n-\ell}(b_0 + \ell a_2; t_1) \end{pmatrix}.$$

Then, the system of the equations (2.7), (2.8) and (2.9) can be written in matrix form as $\psi_{n-\ell+1}(t_1) = C_\ell(t_1)\psi_{n-\ell}(t_1)$ $(\ell = 1, ..., n)$, and $\psi_0(t_1) = \mathbf{1}_{(n+1)} = (1, 1, ..., 1)'$, where, $\mathbf{1}_{(n+1)}$ denotes the $(n+1)\times 1$ column vector whose components are all unity and $C_\ell(t_1)$ denotes the $\ell \times (\ell+1)$ matrix whose (i,j)th component is given by,

$$(2.10)\quad c_{ij}(\ell; t_1) = \begin{cases} \frac{\alpha_{01}+(\ell-i)a_{11}+(i-1)a_{21}}{\alpha_{01}+\alpha_{02}+(\ell-i)(a_{11}+a_{12})+(i-1)(a_{21}+a_{22})}, & j = i, i = 1, ..., \ell, \\ \frac{\alpha_{02}+(\ell-i)a_{12}+(i-1)a_{22}}{\alpha_{01}+\alpha_{02}+(\ell-i)(a_{11}+a_{12})+(i-1)(a_{21}+a_{22})} t_1, & j = i+1, i = 1, ..., \ell, \\ 0, & \text{otherwise}. \end{cases}$$

Proposition 2.1 *The probability generating function $\psi_n(b_0; t_1)$, the exact distribution of Y_n and its expected value are given by*

$$\psi_n(b_0; t_1) = C_1(t_1)C_2(t_1)\cdots C_n(t_1)\mathbf{1}_{(n+1)} = \prod_{i=1}^{n} C_i(t_1)\mathbf{1}_{(n+1)},$$

$$P(Y_n = y) = \sum_{1\le n_1<\cdots<n_y\le n} C_1(0)\cdots \dot{C}_{n_1}(0)\cdots \dot{C}_{n_y}(0)\cdots C_n(0)\mathbf{1}_{(n+1)},$$

$$E[Y_n; b_0] = \sum_{i=1}^{n} C_1(1)\cdots \dot{C}_i(1)\cdots C_n(1)\mathbf{1}_{(n+1)},$$

$$\textit{where, } \dot{C}_k(t_1) = \frac{dC_k(t_1)}{dt_1} = \left(\frac{dc_{ij}(k; t_1)}{dt_1}\right).$$

□

In a similar way, under an initial urn composition $b_0 = (\alpha_{01}, \alpha_{02}, ..., \alpha_{0m})$, we can solve the system of the equations in Theorem 2.1 by virtue of their linearity and obtain the p.g.f.. However, we do not write it due to lack of space.

Remark 1 *In this Pólya urn model, Kotz, Mahmoud and Robert (2000) derived the exact distribution of Y_n by another approach, and derived the recurrence relation for the expected value. They also reported that the expected value can be derived from the recurrence relation in a case that the constraint is imposed, whereas the expected value can not be derived from it in a case that the constraint is not imposed. Then, we present a useful recurrence for the expected values, as will be shown later.*

2.3 The recurrences for the expected values

In this subsection, we present a method for the exact analysis, which are very simple and suitable for computation by computer algebra systems. It is a recurrence for obtaining the expected values for all n.

Theorem 2.1 (The Pólya urn model containing m different labels)
The expected values of $X_n^{(i)}$ $(i = 0, 1, ...r)$, $E[X_n^{(i)}; \boldsymbol{b}]$ *say, satisfy the recurrences;*

$$(2.11) \quad E[X_n^{(i)}; \boldsymbol{b}] = \sum_{j=1}^{m} \frac{\alpha_j}{|\boldsymbol{b}|} \left(I_{B_i}(j) + E[X_{n-1}^{(i)}; \boldsymbol{b} + \boldsymbol{a}_j] \right), \quad n \geq 1, \ i = 1, ..., r,$$

$$(2.12) \quad E[X_0^{(i)}; \boldsymbol{b}] = 0, \qquad i = 1, ..., r.$$

Proof. It is easy to check the equation (2.12). The equation (2.11) is obtained by differentiating both sides of the equation (2.1) with respect to t_i $(i = 1, ..., r)$ and then setting $t_1 = \cdots = t_r = 1$. The proof is completed. □

As a special case, for $m = 2$, we consider the Pólya urn model containing 2 different labels treated in Section 2.2. Then, from Theorem 2.1, we have the following Corollary 2.2.

Corollary 2.2 (The Pólya urn model containing 2 different labels)
The expected value of Y_n, $E[Y_n; \boldsymbol{b}]$ *say, satisfies the recurrence;*

$$(2.13) \qquad E[Y_n; \boldsymbol{b}] = \frac{\alpha_1}{|\boldsymbol{b}|} E[Y_{n-1}; \boldsymbol{b} + \boldsymbol{a}_1] + \frac{\alpha_2}{|\boldsymbol{b}|} E[Y_{n-1}; \boldsymbol{b} + \boldsymbol{a}_2] + \frac{\alpha_2}{|\boldsymbol{b}|}, \quad n \geq 1,$$

$$(2.14) \qquad E[Y_0; \boldsymbol{b}] = 0.$$

□

3 Numerical examples

In this section, we illustrate how to obtain the distributions and the expected values by using computer algebra systems.

Example 4.1 : The Pólya urn model containing 4 different labels
Assume that $\boldsymbol{b}_0 = (1, 2, 1, 2)$, $B = \{1, 2, 3, 4\}$, $B_1 = \{2, 4\}$, $B_2 = \{3, 4\}$ and $A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$. Let $X_n^{(i)} = \sum_{j=1}^{n} I_{B_i}(Z_j)$ $(i = 1, 2)$. For $n = 3$, the p.g.f. is

$$\begin{aligned}\phi_3(\boldsymbol{b}_0; t_1, t_2) = {} & \frac{1}{108} + \frac{73}{3564} t_1 + \frac{73}{2376} t_2 + \frac{269}{7920} {t_1}^2 + \frac{3209}{35640} t_1 t_2 + \frac{269}{7920} {t_2}^2 \\ & + \frac{1}{20} {t_1}^3 + \frac{3361}{23760} {t_1}^2 t_2 + \frac{3191}{35640} t_1 {t_2}^2 + \frac{1}{80} {t_2}^3 + \frac{269}{1980} {t_1}^3 t_2 + \frac{4951}{35640} {t_1}^2 {t_2}^2 \\ & + \frac{269}{11880} t_1 {t_2}^3 + \frac{73}{594} {t_1}^3 {t_2}^2 + \frac{73}{2376} {t_1}^2 {t_2}^3 + \frac{1}{27} {t_1}^3 {t_2}^3.\end{aligned}$$

Table 1. The exact joint probability function of $(X_3^{(1)}, X_3^{(2)})$, given $\boldsymbol{b}_0 = (1,2,1,2)$.

	$X_3^{(1)}=0$	$X_3^{(1)}=1$	$X_3^{(1)}=2$	$X_3^{(1)}=3$
$X_3^{(2)}=0$	0.009259	0.020482	0.033965	0.05
$X_3^{(2)}=1$	0.030724	0.090039	0.141456	0.135859
$X_3^{(2)}=2$	0.033965	0.089534	0.138917	0.122896
$X_3^{(2)}=3$	0.0125	0.022643	0.030724	0.037037

For $n = 10$, we give Fig.1, which is the three-dimensional plot of the exact joint probability function of $(X_{10}^{(1)}, X_{10}^{(2)})$, given $\boldsymbol{b}_0 = (1,2,1,2)$ and the addition matrix A.

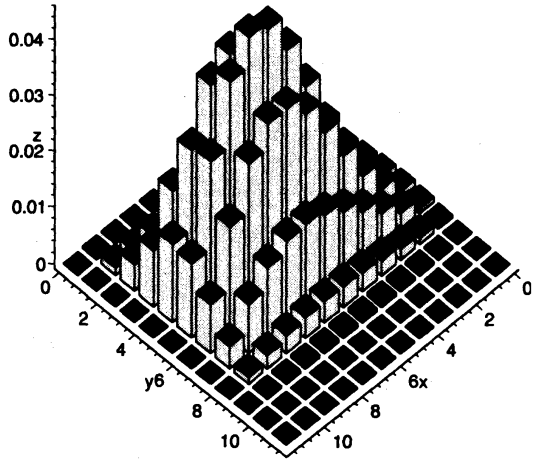

Fig.1. The exact joint probability function of $(X_{10}^{(1)}, X_{10}^{(2)})$ in the Example 4.1, given $\boldsymbol{b}_0 = (1,2,1,2)$ and the addition matrix A.

Marshall (1990) discussed this model in case that the addition matrix is the identity matrix. So far as we know, it was first proposed by Kaiser and Stefansky (1972).

Example 4.2 : The Pólya urn model containing 2 different labels

Assume that $\boldsymbol{b}_0 = (2,3)$, $B = \{1,2\}$, $B_1 = \{2\}$ and $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Let $Y_n = \sum_{j=1}^{n} I_{B_1}(Z_j)$. For $n = 10$, the p.g.f. and the expected value are, respectively,

$$\psi_{10}(\boldsymbol{b}_0; t_1) = \frac{1}{91} + \frac{125291}{3153150} t_1 + \frac{4404557}{50450400} {t_1}^2 + \frac{52734593}{367567200} {t_1}^3 + \frac{8659858873}{46313467200} {t_1}^4$$

$$+ \frac{985104707}{5028319296} {t_1}^5 + \frac{195631373}{1197218880} {t_1}^6 + \frac{8913571}{84651840} {t_1}^7 + \frac{16000}{323323} {t_1}^8$$

$$+ \frac{10240}{676039} {t_1}^9 + \frac{1536}{676039} {t_1}^{10},$$

$$E[Y_{10}; \boldsymbol{b}_0] = \dot{\psi}_{10}(\boldsymbol{b}_0; 1) = \frac{11750459755829}{2529873145800} = 4.644683381.$$

We give Fig.2, which is the two-dimensional plot of the exact expected values of Y_n, given three initial urn compositions $\boldsymbol{b}_0 = (1,1), (2,3), (5,1)$ and the addition matrix A.

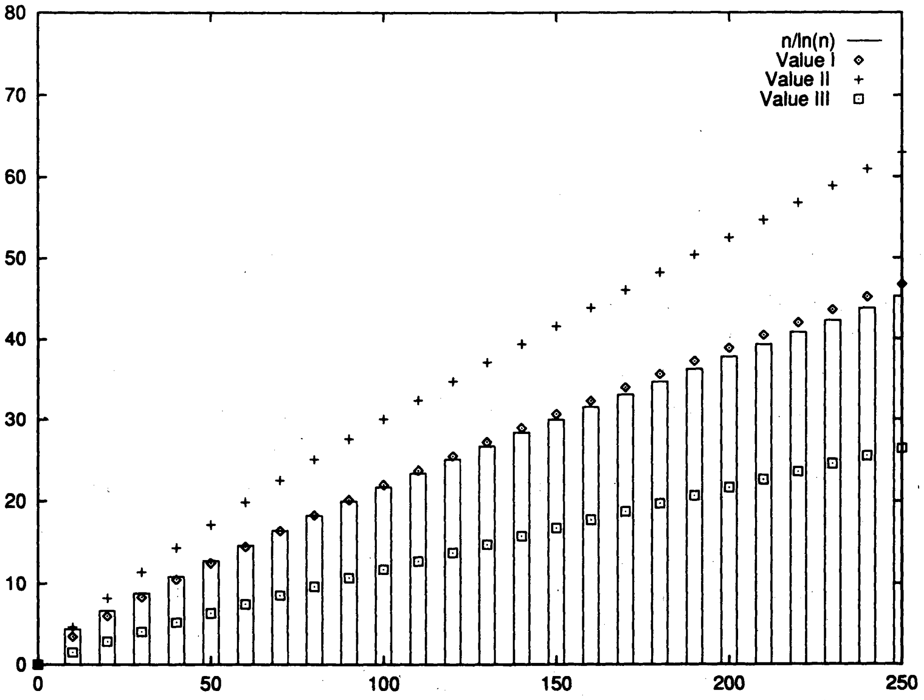

Fig.2. The exact expected values of Y_n, given three initial urn compositions $\boldsymbol{b}_0 = (1,1), (2,3), (5,1)$ and the addition matrix A in Example 4.2. Value I, II, III are, respectively, the values given initial urn compositions $\boldsymbol{b}_0 = (1,1), (2,3), (5,1)$.

Remark 2 *In this example, Kotz, Mahmoud and Robert (2000) suggested that the fixed values of the initial condition will be asymptotically negligible with regard to Y_n for large n and $E[Y_n] \sim n/\ln n$, as $n \to \infty$. By calculating the exact expected values of Y_n given three initial conditions, we observe that their values depend on the initial conditions when n is small.*

However, when n comes to 250, it seems that the exact expected values of Y_n still heavily depend on their initial conditions. Therefore, we think that the exact analysis is important.

REFERENCES

[1] 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.

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

[3] Aki, S. and Hirano, K. (1993), Discrete distributions related to succession events in a two-state Markov chain, *Statistical Sciences and Data Analysis; Proceedings of the Third Pacific Area Statistical Conference* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467–474, VSP International Science Publishers, Zeist.

[4] Aki, S. and Hirano, K. (1999), Sooner and later waiting time problems for runs in Markov dependent bivariate trials, *Ann. Inst. Statist. Math.*,**51**, 17–29.

[5] Bagchi, A. and Pal, A. (1985), Asymptotic normality in the generalized Pólya-Eggenberger urn model with application to computer data structures, *SIAM J. Algebraic Discrete Methods*, **6**, 394–405.

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

[7] Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd ed., Wiley, New York.

[8] Friedman, B. (1949), A simple urn model, *Comm. Pure Appl. Math.*,**2**, 59–70.

[9] Gouet, R. (1989), A martingale approach to strong convergence in a generalized Pólya–Eggenberger urn model, *Statist. Probab. Lett.*,**8**, 225–228.

[10] Gouet, R. (1993), Martingale functional central limit theorems for a generalized Pólya urn, *Ann. Probab.*,**21**, 1624–1639.

[11] Inoue, K. and Aki, S. (2000), Generalized waiting time problems associated with patterns in Pólya's urn scheme, *Research Report on Statistics*, **50**, Osaka University.

[12] Johnson, N. L. and Kotz, S. (1977), *Urn Models and Their Applications*, Wiley, New York.

[13] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1997), *Discrete Multivariate Distributions*, Wiley, New York.

[14] Kaiser, H. F. and Stefansky, W. (1972), A Pólya distribution for teaching, *Amer. Statist.*, **26**, 40–43.

[15] Kocherlakota, S. (1989), A note on the bivariate binomial distribution, *Statist. Probab. Lett.*,**8**, 21–24.

[16] Kotz, S., Mahmoud, H. and Robert, P. (2000), On generalized Pólya urn models, *Statist. Probab. Lett.*,**49**, 163–173.

[17] Mahmoud, H. and Smythe, R. T. (1991), On the distribution of leaves in rooted subtrees of recursive trees, *Ann. Appl. Probab.*, **1**, 406–418.

[18] Marshall, A. W. and Olkin, I. (1985), A family of bivariate distributions generated by the bivariate Bernoulli distribution, *J. Amer. Statist. Assoc.*, **80**, 332–338.

[19] Marshall, A. W. and Olkin, I. (1990), Bivariate distributions generated from Pólya-Eggenberger urn models, *J. Multivariate Anal.*, **35**, 48–65.

[20] Marshall, A. W. and Olkin, I. (1993), Bivariate life distributions from Pólya's urn model for contagion, *J. Appl. Prob.*, **30**, 497–508.

[21] Smythe, R. (1996), Central limit theorems for urn models, *Stochastic Process. Appl.*, **65**, 115–137.