Ann. Inst. Statist. Math.
Vol. 54, No. 3, 681–688 (2002)

GENERALIZED WAITING TIME PROBLEMS ASSOCIATED WITH PATTERN IN POLYA'S URN SCHEME

KIYOSHI INOUE* AND SIGEO AKI

Department of Informatics and Mathematical Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan

(Received July 4, 2000; revised February 21, 2001)

Abstract. Let $X_1, X_2, \ldots$ be a sequence obtained by Polya's urn scheme. We consider a waiting time problem for the first occurrence of a pattern in the sequence $X_1, X_2, \ldots$, which is generalized by a notion "score". The main part of our results is derived by the method of generalized probability generating functions. In Polya's urn scheme, the system of equations is composed of the infinite conditional probability generating functions, which can not be solved. Then, we present a new methodology to obtain the truncated probability generating function in a series up to an arbitrary order from the system of infinite equations. Numerical examples are also given in order to illustrate the feasibility of our results. Our results in this paper are not only new but also a first attempt to treat the system of infinite equations.

Key words and phrases: Polya's urn scheme, pattern, generalized probability generating functions, conditional probability generating functions.

1. Introduction

Exact distributions on runs and patterns in independent trials have been studied for a long time. Distribution theory of runs and patterns has been developed in various situations by many authors, with applications in many areas (see Shmueli and Cohen (2000)), since Ebneshahrashoob and Sobel (1990) solved sooner and later problems for success and failure runs. Furthermore, many problems are treated when the condition that the random variables are independent and identically distributed is widely relaxed. The sooner and later waiting time problems between a success run of length k and a failure run of length r in the first order Markov dependent trials are studied by Aki and Hirano (1993). The problems are extended by Aki *et al.* (1996) in the higher order Markov dependent trials. Uchida (1998) studied some waiting time problems for the patterns in the higher order Markov dependent trials.

There are two standard approaches to these problems. One is a finite Markov chain imbedding technique introduced by Fu and Koutras (1994). This method provides a unified procedure for the distribution of runs, scans and patterns. By this approach, Fu (1996) studied the exact and joint distributions of the runs and patterns in a sequence of multi-state trials. In case of a finite state space, a large number of studies have been made, however, nobody has ever tried to study the case of an infinite state space. The other is to solve a system of equations of conditional probability generating function

*Now at The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan.

(p.g.f.)'s. Then, some characteristics such as probability function and moments are derived from an expansion of the solution. Here, the system of equations is derived by considering the condition of one-step ahead from every condition. Since the number of the conditions are finite until the trials are finished, the system of equations is composed of the finite conditional p.g.f.'s, so that we can solve it. By this approach, Uchida (1998) studied some waiting time problems for the patterns under a certain assumption. However, it is a disadvantage of this method that the system of equations can not be solved when the system is composed of infinite conditional p.g.f.'s. So far as the authors know, the waiting time problems related to such a system of infinite equations have never been examined and no studies have ever tried to treat the system of infinite equations.

The aim of this paper is to propose a method to obtain the probability function from the system of infinite equations. Our methodology is based on the idea of using the conditional p.g.f.'s recursively within the finite procedures. Then, we can obtain the truncated p.g.f. in a series up to an arbitrary order. Of course, our method is useful also for the system of equations composed of finite conditional p.g.f.'s. We consider the generalized waiting time problem for the first occurrence of a pattern in the sequence $X_1, X_2, \ldots$ obtained by Polya's urn scheme. Polya's urn scheme is an interesting process whose conditions of the balls change as the trials, so that the system of equations is composed of the infinite conditional p.g.f.'s.

Let us consider Polya's urn scheme (See Feller (1968)). Suppose that we have α_0 balls labeled "0", α_1 balls labeled "1", ..., α_m balls labeled "m" in an urn. We draw a ball at random and, before drawing the next ball, we replace the ball drawn, adding also c balls of the same label. This procedure is repeated. We denote the number of the balls by $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_m)$ and the amount of the balls by $|\boldsymbol{\alpha}| = \alpha_0 + \alpha_1 + \cdots + \alpha_m$. Let $\boldsymbol{e}_j = (0, 0, \ldots, 0, 1, 0, \ldots, 0)_{1\times(m+1)}$ (the $(j+1)$-th element is 1, and the other elements are all 0, $j = 0, 1, \ldots, m$). If the ball labeled j is drawn, the number of the balls changes to $\boldsymbol{\alpha} + c\boldsymbol{e}_j$.

Let $X_1, X_2, \ldots$ be a sequence of random variables obtained by Polya's urn scheme, which take values in a finite set $\mathcal{B} = \{0, 1, 2, \ldots, m\}$ and let $T = (a_1, a_2, \ldots, a_k)$ be a pattern whose elements are integers in $\mathcal{B}$. In this paper, we shall use the term "pattern" to refer to the finite sequence.

We generalize the waiting time problem for the first occurrence of the pattern by introducing a "score". Suppose that the balls labeled "j" have a score "r_j", $r_j \in \mathcal{N} = \{1, 2, \ldots\}$, $(j = 0, 1, \ldots, m)$. We consider the distribution of the total score until the occurrence of the pattern in the sequence $X_1, X_2, \ldots$ for the first time.

We consider such a generalized waiting time problem for the pattern in Polya's urn scheme and give a general method to obtain the probability function.

Remark that if all the scores r_j are equal to 1, this problem corresponds to the waiting time problem for the occurrence of the pattern in the sequence $X_1, X_2, \ldots$ for the first time.

In Section 2, we consider the distribution of the total score until the first occurrence of the pattern in the sequence $X_1, X_2, \ldots$. We derive the system of equations of conditional p.g.f.'s, which is composed of the infinite equations. We present the method to obtain the truncated p.g.f. in a series up to the arbitrary order. In this method, the score plays an important role.

The results in this paper are not only general and new but also available to numerical and symbolic calculations by using computer algebra systems. In Section 3, two numerical examples are given to illustrate the method and results developed in the pre-

vious section. We can obtain the truncated p.g.f. of the waiting time distribution for the pattern by using computer algebra systems.

2. Distributions of total scores

Let $X_1, X_2, \ldots$ be a sequence of random variables obtained by Polya's urn scheme, which take values in a finite set $\mathcal{B} = \{0, 1, 2, \ldots, m\}$. Let $T = (a_1, a_2, \ldots, a_k)$ be a pattern, which is a finite sequence whose elements are integers in $\mathcal{B}$. Let $T_0 = \emptyset, T_k = T, T_i = (a_1, a_2, \ldots, a_i)(i = 1, 2, \ldots, k)$, then, we define a set $\mathcal{P}(T)$ by $\mathcal{P}(T) = \{T_0, T_1, \ldots, T_k\}$. Let P_{ij} be the longest pattern among $\{(a_1, \ldots, a_i, j), (a_2, \ldots, a_i, j), \ldots, (j), \emptyset\} \cap \mathcal{P}(T)$. Let $f : (\mathcal{P}(T) \setminus \{T\}) \times \mathcal{B} \to \mathcal{P}(T)$ be a mapping defined by $f(T_i, j) = P_{ij}$.

Example 2.1. Assume that $\mathcal{B} = \{0, 1, 2\}$, $T = (0, 0, 0, 1)$ and $\mathcal{P}(T) = \{T_0, T_1, T_3, T_4\}$. Suppose that we have currently $T_3 = (0, 0, 0)$. After the next draw, if the ball labeled 0 is drawn, we have $f(T_3, 0) = T_3$, if the ball labeled 1 is drawn, we have $f(T_3, 1) = T_4$, if the ball labeled 2 is drawn, we have $f(T_3, 2) = T_0$.

Let $r(\cdot)$ be a function which maps from $\mathcal{B}$ to $\mathcal{N}$. To simplify the notation, denote $r(i)$ by r_i, and we call this non-negative integer r_i "score". By considering the following example, we will find it natural to introduce the score.

Example 2.2. Suppose that a machine is inspected once a day and its condition is recorded, whose condition is classified into 5 levels (say $1, 2, 3, 4, 5$). If "i" $(i = 1, \ldots, 4)$ is observed, then, the machine costs $r_i (i = 1, \ldots, 4)$ for the each repair. If "5" is observed, the machine must be stopped by the replacement.

In Example 2.2, the cost $r_i (i = 1, \ldots, 4)$ for the each repair corresponds to the score and the expenses of the repair until the replacement corresponds to the total score.

Suppose that we have $T_i (i = 0, \ldots, k)$, the balls $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, \ldots, \alpha_m)$ and the score n and the first T has not yet occurred. Then, we denote by $S(T_i, \boldsymbol{\alpha}, n)$ the score obtained from this time until the pattern T occurs for the first time. Let W_T be the waiting time for the first pattern T. The total score $S(T_0, \boldsymbol{\alpha}_0, 0)$ until the pattern T occurs for the first time is expressed by

$$S(T_0, \boldsymbol{\alpha}_0, 0) = \sum_{i=1}^{W_T} \sum_{j=0}^{m} r(j, X_i), \quad \text{where,} \quad r(j, X_i) = \begin{cases} r_j & \text{if } X_i = j, \\ 0 & \text{otherwise.} \end{cases}$$

We consider the distribution of the total score until the first occurrence of the pattern T. Let $\Phi(t)$ be the p.g.f. of the distribution of the total score until the pattern T occurs for the first time in $X_1, X_2, \ldots$ under the initial balls $\boldsymbol{\alpha}_0$ in the urn. Suppose that we have currently $T_i (i = 0, 1, \ldots, k)$, the score n and the balls $\boldsymbol{\alpha}$ in the urn. Then, we denote by $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k)$ the conditional p.g.f. of the distribution of the total score from this time until the pattern T occurs for the first time. Easily, we see that $\Phi(t) = \phi(T_0, \boldsymbol{\alpha}_0, 0; t)$. $\Phi(t)$, $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k)$ are defined by

$$\Phi(t) = E[t^{S(T_0, \boldsymbol{\alpha}_0, 0)}], \tag{2.1}$$

$$\phi(T_i, \boldsymbol{\alpha}, n; t) = E[t^{S(T_i, \boldsymbol{\alpha}, n)}], \quad i = 0, 1, \ldots, k. \tag{2.2}$$

From the definitions of $\Phi(t)$ and $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k)$, we can derive the following system of infinite equations:

THEOREM 2.1. *The p.g.f.* $\Phi(t)$ *and the conditional p.g.f.'s* $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k)$ *satisfy the following system of equations*:

$$\Phi(t) = \sum_{j=0}^{m} \frac{\alpha_{0j}}{|\boldsymbol{\alpha}_0|} t^{r_j} \phi(f(T_0, j), \boldsymbol{\alpha}_0 + c\boldsymbol{e}_j, r_j; t), \tag{2.3}$$

$$\phi(T_i, \boldsymbol{\alpha}, n; t) = \sum_{j=0}^{m} \frac{\alpha_j}{|\boldsymbol{\alpha}|} t^{r_j} \phi(f(T_i, j), \boldsymbol{\alpha} + c\boldsymbol{e}_j, n + r_j; t), \tag{2.4}$$

$$i = 0, 1, 2, \ldots, k-2,$$

$$\phi(T_{k-1}, \boldsymbol{\alpha}, n; t) = \sum_{j=0}^{m} \frac{\alpha_j}{|\boldsymbol{\alpha}|} t^{r_j} \phi(f(T_{k-1}, j), \boldsymbol{\alpha} + c\boldsymbol{e}_j, n + r_j; t), \tag{2.5}$$

$$\phi(T_k, \boldsymbol{\alpha}, n; t) = 1, \quad \textit{where,} \quad \Phi(t) = \phi(T_0, \boldsymbol{\alpha}_0, 0; t). \tag{2.6}$$

PROOF. It is easy to see that $\phi(T_k, \boldsymbol{\alpha}, n; t) = 1$ and $\Phi(t) = \phi(T_0, \boldsymbol{\alpha}_0, 0; t)$ by the definition of p.g.f. Let τ be the total score until the first occurrence of the pattern T and let n be the arbitrary score. Recall that $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k-1)$ is the p.g.f. of the conditional distribution of the random variable $\tau - n$ given that $T_i (i = 0, 1, \ldots, k-1)$ has just occurred with the score n. We note that $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k-1)$ depends on the event that we have currently $T_i (i = 0, 1, \ldots, k-1)$ and the balls $\boldsymbol{\alpha}$ in the urn. Therefore, by considering the condition of one-step ahead from every condition, we see that $\Phi(t)$ and $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, 1, \ldots, k)$ satisfy the above equations. The proof is completed. □

Example 2.3. *Waiting time for* $T = (1)$. Assume that $\mathcal{B} = \{0, 1\}$, $r_0 = r_1 = 1$, $T_0 = \emptyset$ and $T_1 = (1)$. $\boldsymbol{e}_0 = (1, 0)$, $\boldsymbol{e}_1 = (0, 1)$. Then, the system of the equations is

$$\Phi(t) = \frac{\alpha_{00}}{|\boldsymbol{\alpha}_0|} t\phi(T_0, \boldsymbol{\alpha}_0 + c\boldsymbol{e}_0, 1; t) + \frac{\alpha_{01}}{|\boldsymbol{\alpha}_0|} t\phi(T_1, \boldsymbol{\alpha}_0 + c\boldsymbol{e}_1, 1; t), \tag{2.7}$$

$$\phi(T_0, \boldsymbol{\alpha}, n; t) = \frac{\alpha_0}{|\boldsymbol{\alpha}|} t\phi(T_0, \boldsymbol{\alpha} + c\boldsymbol{e}_0, n + 1; t) + \frac{\alpha_1}{|\boldsymbol{\alpha}|} t\phi(T_1, \boldsymbol{\alpha} + c\boldsymbol{e}_1, n + 1; t), \tag{2.8}$$

$$\phi(T_1, \boldsymbol{\alpha}, n; t) = 1, \quad \text{where,} \quad \Phi(t) = \phi(T_0, \boldsymbol{\alpha}_0, 0; t). \tag{2.9}$$

We will show that the truncated p.g.f. of $\Phi(t)$, say $\hat{\Phi}(t)$, is obtained in a polynomial form of t by using the equations (2.8) and (2.9) for the right-hand side of the equation (2.7) recursively. The idea of truncation is also illustrated.

From the equation (2.8),

$$\phi(T_0, \boldsymbol{\alpha}_0 + c\boldsymbol{e}_0, 1; t) = \frac{\alpha_{00} + c}{|\boldsymbol{\alpha}_0 + c\boldsymbol{e}_0|} t\phi(T_0, \boldsymbol{\alpha}_0 + 2c\boldsymbol{e}_0, 2; t) + \frac{\alpha_{01}}{|\boldsymbol{\alpha}_0 + c\boldsymbol{e}_0|} t\phi(T_1, \boldsymbol{\alpha}_0 + c(\boldsymbol{e}_0 + \boldsymbol{e}_1), 2; t). \tag{2.10}$$

By substituting (2.10) into the right-hand side of (2.7), we have

$$\Phi(t) = \frac{\alpha_{00}}{|\boldsymbol{\alpha}_0|} t \left(\frac{\alpha_{00} + c}{|\boldsymbol{\alpha}_0 + c\boldsymbol{e}_0|} t\phi(T_0, \boldsymbol{\alpha}_0 + 2c\boldsymbol{e}_0, 2; t) \right. \tag{2.11}$$

$$+\frac{\alpha_{01}}{|\boldsymbol{\alpha}_0+c\boldsymbol{e}_0|}t\phi(T_1,\boldsymbol{\alpha}_0+c(\boldsymbol{e}_0+\boldsymbol{e}_1),2;t)\Big)$$
$$+\frac{\alpha_{01}}{|\boldsymbol{\alpha}_0|}t\phi(T_1,\boldsymbol{\alpha}_0+c\boldsymbol{e}_1,1;t).$$

If we need the truncated p.g.f. of $\Phi(t)$ in a polynomial form of t up to the first order, we should set $\phi(T_0,\boldsymbol{\alpha}_0+2c\boldsymbol{e}_0,2;t)=\phi(T_1,\boldsymbol{\alpha}_0+c(\boldsymbol{e}_0+\boldsymbol{e}_1),2;t)=0$ and use the equation (2.9) in the right-hand side of (2.11). Then, we obtain the truncated generating function of $\Phi(t)$, say $\hat{\Phi}(t)$;

$$\hat{\Phi}(t)=\frac{\alpha_{01}}{|\boldsymbol{\alpha}_0|}t.$$

Thus, the above substitution leads to the truncated generating function $\hat{\Phi}(t)$. If we need the truncated p.g.f. of $\Phi(t)$ in a polynomial form of t up to the second order, again, we substitute the equation (2.8) into the term $\phi(T_0,\boldsymbol{\alpha}_0+2c\boldsymbol{e}_0,2;t)$ in the right-hand side of (2.11), respectively. Setting $\phi(T_0,\boldsymbol{\alpha}_0+3c\boldsymbol{e}_0,3;t)=\phi(T_1,\boldsymbol{\alpha}_0+2c\boldsymbol{e}_0+c\boldsymbol{e}_1,3;t)=0$ and using the equation (2.9), we have the truncated p.g.f. which is truncated in a polynomial form of t up to the second order. If we need truncated p.g.f. of $\Phi(t)$ in a polynomial form of t up to n_0-th order, we should continue this procedure until the score in the conditional p.g.f. $\phi(T_0,\boldsymbol{\alpha},n;t)$ in the right-hand side of the equation (2.11) becomes greater than n_0. Then, we let all the conditional p.g.f.'s whose scores are greater than n_0 be equal to zero and use the equation (2.9) in the right-hand side of the equation (2.11). Notice that we have the polynomial form up to the n_0-th order and the order of the vanished terms is greater than n_0. Consequently, we obtain the truncated p.g.f. of $\Phi(t)$ in a series of t up to the n_0-th order.

THEOREM 2.2. *For any positive integer n_0, the following system of equations leads to the truncated p.g.f. of $\Phi(t)$, $\hat{\Phi}(t)$ say, which is expanded in a power series of t up to the n_0-th order.*

$$\hat{\Phi}(t)=\sum_{j=0}^{m}\frac{\alpha_{0j}}{|\boldsymbol{\alpha}_0|}t^{r_j}\hat{\phi}(f(T_0,j),\boldsymbol{\alpha}_0+c\boldsymbol{e}_j,r_j;t), \tag{2.12}$$

$$\hat{\phi}(T_i,\boldsymbol{\alpha},n;t)=\sum_{j=0}^{m}\frac{\alpha_j}{|\boldsymbol{\alpha}|}t^{r_j}\hat{\phi}(f(T_i,j),\boldsymbol{\alpha}+c\boldsymbol{e}_j,n+r_j;t), \tag{2.13}$$
$$i=0,1,\ldots,k-2,\ n\le n_0,$$

$$\hat{\phi}(T_{k-1},\boldsymbol{\alpha},n;t)=\sum_{j=0}^{m}\frac{\alpha_j}{|\boldsymbol{\alpha}|}t^{r_j}\hat{\phi}(f(T_{k-1},j),\boldsymbol{\alpha}+c\boldsymbol{e}_j,n+r_j;t),\qquad n\le n_0, \tag{2.14}$$

$$\hat{\phi}(T_i,\boldsymbol{\alpha},n;t)=0,\qquad i=0,1,\ldots,k,\ n>n_0, \tag{2.15}$$

$$\hat{\phi}(T_k,\boldsymbol{\alpha},n;t)=1,\qquad n\le n_0. \tag{2.16}$$

PROOF. In the equations (2.4) and (2.5), we see that the score in the right-hand side is greater than the one of the left-hand side. Notice that $\phi(T_i,\boldsymbol{\alpha},n;t)$ $(i=0,\ldots,k-1)$ is expressed by the conditional p.g.f.'s with the larger scores. By using the equations (2.4) and (2.5) for obtaining every term of the right-hand side of equation (2.3) recursively, the p.g.f. $\Phi(t)$ is expressed by the conditional p.g.f.'s with the larger score. If

we need truncated p.g.f. of $\Phi(t)$ in a series of t up to n_0-th order, we should continue this procedure until all the scores in the conditional p.g.f.'s $\phi(T_i, \boldsymbol{\alpha}, n; t)$ $(i = 0, \ldots, k-1)$ in the right-hand side of the equation (2.3) become greater than n_0. Then, we let all the conditional p.g.f.'s whose scores are greater than n_0 be equal to zero and use the equation (2.6) in the right-hand side of the equation (2.3). We should notice that the polynomial terms are up to the n_0-th order and the order of the vanished terms is greater than n_0. We obtain the truncated p.g.f. of $\Phi(t)$ in a series of t up to the n_0-th order. Consequently, we can obtain the truncated p.g.f. of $\Phi(t)$ in a series of t up to all order, since a positive number n_0 is arbitrary. The proof is completed. □

Remark 1. As mentioned previously, if all the score $r_i (i = 0, 1, \ldots, m)$ are equal to 1, the p.g.f. $\Phi(t)$ corresponds to the p.g.f. of the waiting time distribution for the first occurrence of the pattern. Our method in Theorem 2.2 can apply to the waiting time problem as well.

3. Numerical examples

In this section, we give some examples and illustrate the waiting time distributions. For numerical computation, Theorem 2.2 is convenient, since the system of equations (2.12), (2.13), (2.14), (2.15) and (2.16) is just the algorithm for computer algebra systems to obtain the truncated p.g.f. of the waiting time distribution for the pattern. We derive the truncated p.g.f. of the waiting time distribution for the pattern by using computer algebra system.

Example 3.1. Assume that $T = (0, 1, 0)$, $\mathcal{B} = \{0, 1\}$, $\alpha_{00} = \alpha_{01} = 2$, $c = 1$ and $r_0 = 3$, $r_1 = 2$. By using the algorithm given by Theorem 2.2, for $n_0 = 100$, we can get the truncated p.g.f. of $\Phi_T(t)$, $\hat{\Phi}_T(t)$ say. Since it is very large, we give Fig. 1 and $\hat{\Phi}_T(t)$ in a series of t up to t^{20} for lack of space.

$$\hat{\Phi}_T(t) = \frac{1}{10}t^8 + \frac{3}{70}t^{10} + \frac{2}{35}t^{11} + \frac{3}{140}t^{12} + \frac{3}{140}t^{13} + \frac{1}{21}t^{14} + \frac{2}{105}t^{15} + \frac{2}{105}t^{16} + \frac{4}{105}t^{17} + \frac{29}{1540}t^{18} + \frac{27}{1540}t^{19} + \frac{47}{1540}t^{20}.$$

Figure 1 is the graph of probability function of waiting time in Example 3.1.

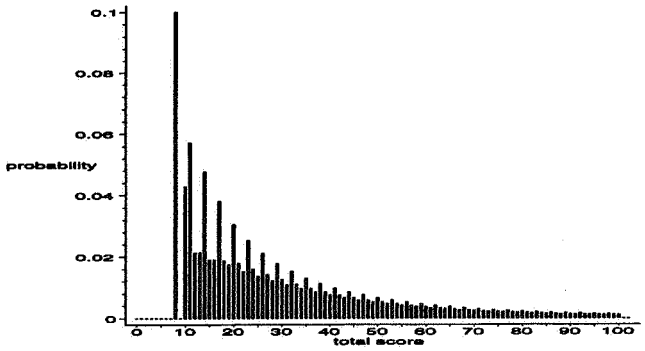

Fig. 1. Probability function of the total score of Example 3.1, given $n_0 = 100$.

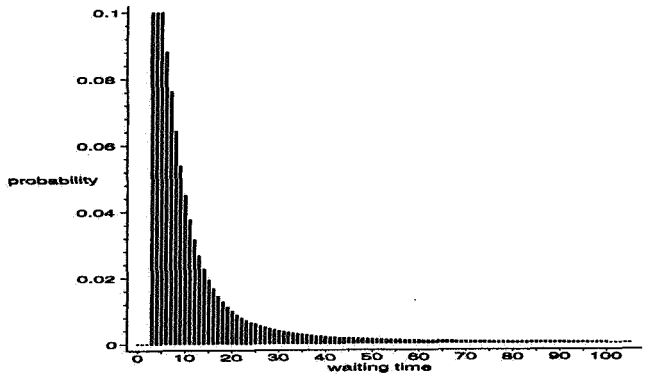

Fig. 2. Probability function of the waiting time of Example 3.2, given $n_0 = 100$.

Example 3.2. Assume that $U = (1, 1, 0)$, $\mathcal{B} = \{0, 1\}$, $\alpha_{00} = \alpha_{01} = 2$, $c = 1$ and $r_0 = r_1 = 1$. By using the algorithm given by Theorem 2.2, for $n_0 = 100$, we can get the truncated p.g.f. of $\Phi_U(t)$, $\hat{\Phi}_U(t)$ say. Since it is very large, we give Fig. 2 and $\hat{\Phi}_U(t)$ in a series of t up to t^{16} for lack of space.

$$\hat{\Phi}_U(t) = \frac{1}{10}t^3 + \frac{1}{10}t^4 + \frac{1}{10}t^5 + \frac{37}{420}t^6 + \frac{8}{105}t^7 + \frac{9}{140}t^8 + \frac{83}{1540}t^9 + \frac{52}{1155}t^{10} + \frac{29}{770}t^{11} + \frac{633}{20020}t^{12} + \frac{229}{8580}t^{13} + \frac{227}{10010}t^{14} + \frac{39631}{2042040}t^{15} + \frac{11391}{680680}t^{16}.$$

Figure 2 is the graph of probability function of waiting time in Example 3.2.

Acknowledgements

We wish to thank the editor and the referees for careful reading of our paper and helpful suggestions which led improved results. Particularly, we would like to thank one of referees for his helpful comments on Theorem 2.1, which led to considerable improvements in the paper.

REFERENCES

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.

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.

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.

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

Fu, J. C. (1996). Distribution theory of runs and patterns associated with a sequence of multi-state trials, *Statist. 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.

Shmueli, G. and Cohen, A. (2000). Run-related probability functions applied to sampling inspection, *Technometrics*, **42**, 188–202.

Uchida, M. (1998). On generating functions of waiting time problems for sequence patterns of discrete random variables, *Ann. Inst. Statist. Math.*, **50**, 655–671.