J. Japan Statist. Soc.
Vol. 38 No. 1 2008 97–105

JOINT DISTRIBUTIONS OF WAITING TIME RANDOM VARIABLES FOR PATTERNS

Sigeo Aki* and Katuomi Hirano**

Exact joint distributions of waiting times for two patterns in a sequence of ℓ-th order time-homogeneous Markov dependent trials are studied, where the patterns are not necessarily assumed to be distinct from each other. We prove that exact joint probability generating functions, which are regarded as expectations of the corresponding random variables, are derived through calculating the conditional expectation based on conditioning by the sooner waiting time and the pattern which comes sooner. We also give illustrative numerical examples in order to demonstrate the performance of our results.

Key words and phrases: Conditional expectation, discrete distribution theory, discrete pattern, generalized probability generating function, waiting time.

1. Introduction

When we are interested in waiting times for two patterns which never occur simultaneously, for example, the waiting times for a success run and a failure run, we usually study the sooner and later waiting times. In this case, we can readily obtain the distribution of the sooner waiting time and the conditional distribution of the later waiting time given the sooner waiting time and which pattern comes sooner, and hence we can derive the joint distribution of the sooner and later waiting times. However, if we can not assume the above condition, it is not obvious to obtain the joint distribution of waiting times for two patterns from the distributions of the sooner and the later waiting times. We propose in this paper a general method for obtaining the joint distribution of the waiting times for two patterns. In our approach, the sooner waiting time also plays an important role.

Let S be a finite set and let $\{X_t\}_{t=1}^{\infty}$ be an ℓ-th order time-homogeneous S-valued Markov chain with $P(X_1 = x_1, \ldots, X_r = x_r) = p_{x_1,\ldots,x_r}$, $r = 1, 2, \ldots, \ell$ and $P(X_{i+\ell} = x_{\ell+1} \mid X_i = x_1, \ldots, X_{i+\ell-1} = x_\ell) = p_{(x_1,\ldots,x_\ell),x_{\ell+1}}$, for $i = 1, 2, \ldots$. Let $\{Y_t\}_{t=-\infty}^{\infty}$ be an ℓ-th order time-homogeneous S-valued Markov chain with $P(Y_{i+\ell} = y_{\ell+1} \mid Y_i = y_1, \ldots, Y_{i+\ell-1} = y_\ell) = p_{(y_1,\ldots,y_\ell),y_{\ell+1}}$, for $i = 1, 2, \ldots$, that is, $\{Y_t\}_{t=-\infty}^{\infty}$ is the two-sided sequence of the ℓ-th order time-homogeneous S-valued Markov dependent trials with the same transition probabilities as $\{X_t\}_{t=1}^{\infty}$. A finite sequence of elements of S is called a pattern. Let $A = (a_1, a_2, \ldots, a_m)$

Accepted October 12, 2007.

*Department of Mathematics, Faculty of Engineering Science, Kansai University, Suita, Osaka 564-8680, Japan.

**The Institute of Statistical Mathematics, Research Organization of Information and Systems, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan (Department of Statistical Science, School of Multidisciplinary Sciences, The Graduate University for Advanced Studies).

and $B = (b_1, b_2, \ldots, b_n)$ be two patterns. W_A and W_B denote the waiting time in $\{X_t\}_{t=1}^{\infty}$ for A and B, respectively. $\overline{W_A}$ and $\overline{W_B}$ denote the waiting time in $\{Y_t\}_{t=1}^{\infty}$ for A and B, respectively, taking the values of $\{Y_t\}_{t=-\infty}^{0}$ into consideration. We assume $m \le n$ without loss of generality and $\ell \le m$ is also assumed for simplicity.

We study the joint distribution of W_A and W_B. Our approach to the joint distribution is a method for obtaining the expectation of $s^{W_A} t^{W_B}$ by using the stepwise smoothing formula of the conditional expectations. The essential part of this approach is to make conditioning by the sooner waiting time between W_A and W_B. Though the distributions of the sooner waiting times for several patterns have been investigated by many authors (see e.g. Fu (1996), Fu and Chang (2002), Balakrishnan and Koutras (2002), Fu and Lou (2003)), there were some constraints on the patterns such that every pattern never be a subpattern of other patterns. When we study only the sooner waiting time, such constraints may be reasonable in order to simplify the final result. When we study joint distributions of some waiting time variates, however, we have to investigate the sooner waiting time without such constraints, because every marginal distribution of the waiting time for one pattern should be derived from the joint distribution.

2. Joint distribution

For patterns $A = (a_1, a_2, \ldots, a_m)$ and $B = (b_1, b_2, \ldots, b_n)$, we define some sets of subpatterns from the left end and from the right end as follows:

$$
\begin{aligned}
SPL(A) &= \{(), (a_1), (a_1, a_2), \ldots, (a_1, a_2, \ldots, a_{m-1}), A\},\\
SPL(B) &= \{(), (b_1), (b_1, b_2), \ldots, (b_1, b_2, \ldots, b_{n-1}), B\},\\
SPL_0(A) &= SPL(A) \setminus \{A\},\\
SPL_0(B) &= SPL(B) \setminus \{B\},\\
SPR(A) &= \{(), (a_m), (a_{m-1}, a_m), \ldots, (a_2, a_3, \ldots, a_m), A\},\\
SPR(B) &= \{(), (b_n), (b_{n-1}, b_n), \ldots, (b_2, b_3, \ldots, b_n), B\},
\end{aligned}
$$

where we are assuming that the empty pattern denoted by () is an element of every set of subpatterns. We denote by $A \odot B$ the longest element in $SPR(A) \cap SPL_0(B)$ and denote by $B \odot A$ the longest element in $SPR(B) \cap SPL_0(A)$. Let $\tau = \min(W_A, W_B)$.

First, we study the case that there exits $j (> m)$ such that $(b_1, b_2, \ldots, b_j) \in SPL_0(B)$ and $a_1 = b_{j-m+1}, a_2 = b_{j-m+2}, \ldots, a_m = b_j$. In this case, j^* denotes the minimum of j which satisfies the above condition. In this case, $\tau = W_A < W_B$ necessarily holds. Let E_1 be the event that $\{(X_{\tau-j^*+1} = b_1, X_{\tau-j^*+2} = b_2, \ldots, X_\tau = b_{j^*})$ occurs$\}$. Then, we can calculate the joint probability generating function (p.g.f.) as follows:

$$
\begin{aligned}
\phi(s,t) &= E[s^{W_A} t^{W_B}]\\
&= E[(st)^{\tau} t^{W_B - \tau}]\\
&= E[(st)^{\tau} t^{W_B - \tau} I_{E_1}] + E[(st)^{\tau} t^{W_B - \tau} I_{E_1^c}]
\end{aligned}
$$

$$
\begin{aligned}
&= E[E[(st)^\tau t^{W_B-\tau} I_{E_1} \mid (\tau, I_{E_1})]] \\
&\quad + E[E[(st)^\tau t^{W_B-\tau} I_{E_1^c} \mid (\tau, I_{E_1^c})]] \\
&= E[(st)^\tau I_{E_1} E[t^{W_B-\tau} \mid (\tau, I_{E_1})]] \\
&\quad + E[(st)^\tau I_{E_1^c} E[t^{W_B-\tau} \mid (\tau, I_{E_1^c})]].
\end{aligned}
$$

Here, since the underlying sequence is time-homogeneous, we can write

$$
\begin{aligned}
&I_{E_1} E[t^{W_B-\tau} \mid (\tau, I_{E_1})] \\
&\qquad = I_{E_1} E[t^{\overline{W_B}} \mid Y_{-j^*+1} = b_1, Y_{-j^*+2} = b_2, \ldots, Y_0 = b_{j^*}] \\
&\qquad \equiv I_{E_1} \psi_B(t; (b_1, \ldots, b_{j^*}), (a_{m-\ell+1}, \ldots, a_m)) \qquad \text{(say)}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&I_{E_1^c} E[t^{W_B-\tau} \mid (\tau, I_{E_1^c})] \\
&\qquad = I_{E_1^c} E[t^{\overline{W_B}} \mid Y_{-m+1} = a_1, Y_{-m+2} = a_2, \ldots, Y_0 = a_m] \\
&\qquad \equiv I_{E_1^c} \psi_B(t; A \odot B, (a_{m-\ell+1}, \ldots, a_m)) \qquad \text{(say)}.
\end{aligned}
$$

Next, we treat the other case, that is, the case where $j(> m)$ does not exist such that $(b_1, b_2, \ldots, b_j) \in SPL_0(B)$ and $a_1 = b_{j-m+1}, a_2 = b_{j-m+2}, \ldots, a_m = b_j$. In this case, it is random which comes sooner between A and B. Therefore, we can calculate the joint p.g.f. as follows:

$$
\begin{aligned}
\phi(s,t) &= E[s^{W_A} t^{W_B}] \\
&= E[(st)^\tau s^{W_A-\tau} t^{W_B-\tau}] \\
&= E[(st)^\tau t^{W_B-\tau} I_{\{\tau=W_A<W_B\}}] \\
&\quad + E[(st)^\tau s^{W_A-\tau} I_{\{\tau=W_B<W_A\}}] \\
&\quad + E[(st)^\tau I_{\{\tau=W_A=W_B\}}] \\
&= E[(st)^\tau I_{\{\tau=W_A<W_B\}} E[t^{W_B-\tau} \mid (\tau, I_{\{\tau=W_A<W_B\}})]] \\
&\quad + E[(st)^\tau I_{\{\tau=W_B<W_A\}} E[s^{W_A-\tau} \mid (\tau, I_{\{\tau=W_B<W_A\}})]] \\
&\quad + E[(st)^\tau I_{\{\tau=W_A=W_B\}}].
\end{aligned}
$$

Here, the value of the conditional expectation $E[t^{W_B-\tau} \mid (\tau, I_{\{\tau=W_A<W_B\}})]$ on $\{\tau = W_A < W_B\}$ and the value of the conditional expectation $E[s^{W_A-\tau} \mid (\tau, I_{\{\tau=W_B<W_A\}})]$ on $\{\tau = W_B < W_A\}$ do not depend on τ. In fact, we see that

$$
\begin{aligned}
&I_{\{\tau=W_A<W_B\}} E[t^{W_B-\tau} \mid (\tau, I_{\{\tau=W_A<W_B\}})] \\
&\qquad = I_{\{\tau=W_A<W_B\}} E[t^{\overline{W_B}} \mid Y_{-m+1} = a_1, Y_{-m+2} = a_2, \ldots, Y_0 = a_m] \\
&\qquad \equiv I_{\{\tau=W_A<W_B\}} \psi_B(t; A \odot B, (a_{m-\ell+1}, \ldots, a_m)) \qquad \text{(say)},
\end{aligned}
$$

and

$$
\begin{aligned}
&I_{\{\tau=W_B<W_A\}} E[s^{W_A-\tau} \mid (\tau, I_{\{\tau=W_B<W_A\}})] \\
&\qquad = I_{\{\tau=W_B<W_A\}} E[s^{\overline{W_A}} \mid Y_{-n+1} = b_1, Y_{-n+2} = b_2, \ldots, Y_0 = b_n] \\
&\qquad \equiv I_{\{\tau=W_B<W_A\}} \psi_A(s; B \odot A, (b_{n-\ell+1}, \ldots, b_n)) \qquad \text{(say)}.
\end{aligned}
$$

Summarizing, we have the following result.

THEOREM 1. *Let* $\{X_t\}_{t=1}^{\infty}$ *be an* ℓ*-th order time-homogeneous* S*-valued Markov chain with the initial probabilities* $P(X_1 = x_1, \dots, X_\ell = x_r) = p_{x_1,\dots,x_r}$, $r = 1, 2, \dots, \ell$ *and the transition probabilities* $P(X_{i+\ell} = x_{\ell+1} \mid X_i = x_1, \dots, X_{i+\ell-1} = x_\ell) = p_{(x_1,\dots,x_\ell),x_{\ell+1}}$, *for* $i = 1, 2, \dots$. *Then, the joint probability generating function of* W_A *and* W_B *is written as*

$$\begin{aligned}\phi(s,t) = {} & E[(st)^\tau I_{\{\tau=W_A<W_B\}\cap E_1}] \cdot \psi_B(t;(b_1,\dots,b_{j^*}),(a_{m-\ell+1},\dots,a_m)) \\ & + E[(st)^\tau I_{\{\tau=W_A<W_B\}\cap E_1^c}] \cdot \psi_B(t; A\odot B,(a_{m-\ell+1},\dots,a_m)) \\ & + E[(st)^\tau I_{\{\tau=W_B<W_A\}}] \cdot \psi_A(s; B\odot A,(b_{n-\ell+1},\dots,b_n)) \\ & + E[(st)^\tau I_{\{\tau=W_A=W_B\}}],\end{aligned}$$

where $\psi_B(t;(b_1,\dots,b_{j^*}),(a_{m-\ell+1},\dots,a_m))$, $\psi_B(t;A\odot B,(a_{m-\ell+1},\dots,a_m))$, *and* $\psi_A(s;B\odot A,(b_{n-\ell+1},\dots,b_n))$ *are defined above.*

Based on the above theorem, we show how to obtain the joint p.g.f. of W_A and W_B. For two patterns $\alpha = (\alpha_1,\dots,\alpha_a)$ and $\beta = (\beta_1,\dots,\beta_b)$, we let $\langle\alpha,\beta\rangle = (\alpha_1,\dots,\alpha_a,\beta_1,\dots,\beta_b)$ be the concatenated pattern. For $1 \le i \le j \le a$, we denote by $[\alpha]_i^j$ the subpattern $(\alpha_i,\alpha_{i+1},\dots,\alpha_j)$ of α. In the following, we explain how to obtain the joint p.g.f. of W_A and W_B.

First, we calculate the generalized p.g.f. of τ with markers (x_1, x, y, z),

$$\begin{aligned}& E[(st)^\tau I_{\{\tau=W_A<W_B\}\cap E_1}]x_1 + E[(st)^\tau I_{\{\tau=W_A<W_B\}\cap E_1^c}]x \\ & \qquad + E[(st)^\tau I_{\{\tau=W_B<W_A\}}]y + E[(st)^\tau I_{\{\tau=W_A=W_B\}}]z.\end{aligned}$$

Here, the idea of the generalized p.g.f. was introduced by Ebneshahrashoob and Sobel (1990). For a nonnegative integer r, S^r denotes the set of all the patterns of length r. In particular, $S^0 = \{()\}$ denotes the set of the empty pattern $()$. In order to derive the generalized p.g.f. of τ, we define the mapping $f : (SPL_0(A) \cup SPL_0(B)) \times (\bigcup_{r=0}^{\ell} S^r) \times S \to (SPL(A) \cup SPL(B)) \times (\bigcup_{r=0}^{\ell} S^r)$, by $f(\alpha,\beta,u) = (f_1(\alpha,u), f_2(\beta,u))$, where $f_1(\alpha,u)$ is the longest element in $SPR(\langle\alpha,u\rangle) \cap (SPL(A) \cup SPL(B))$, and

$$f_2(\beta,u) = \begin{cases} \langle\beta,u\rangle & \text{if the length of } \beta \text{ is less than } \ell, \\ [\langle\beta,u\rangle]_2^{\ell+1} & \text{if the length of } \beta \text{ is equal to } \ell. \end{cases}$$

We denote by

$$p(\beta,u) = \begin{cases} p_{\beta,u} & \text{if the length of } \beta \text{ is equal to } \ell, \\ \dfrac{p_{\langle\beta,u\rangle}}{p_\beta} & \text{if the length of } \beta \text{ is less than } \ell, \end{cases}$$

the conditional probability that "u" comes next given β, where we set $p_{()} = 1$. For every $\alpha \in SPL_0(A) \cup SPL_0(B)$ and $\beta \in \bigcup_{r=0}^{\ell} S^r$, the pair (α,β) is called a consistent pair if at least $[\alpha]_{a-b+1}^a = \beta$ or $[\beta]_{b-a+1}^b = \alpha$ holds, where a and b are the length of α and β, respectively. We denote by $CP((SPL_0(A) \cup SPL_0(B)) \times \bigcup_{r=0}^{\ell} S^r)$ the set of all the consistent pairs $(\alpha, \beta) \in (SPL_0(A) \cup SPL_0(B)) \times \bigcup_{r=0}^{\ell} S^r$. Similarly, we define $CP(SPL_0(A) \times S^\ell)$ and $CP(SPL_0(B) \times S^\ell)$. For every $(\alpha, \beta) \in CP((SPL_0(A) \cup SPL_0(B)) \times \bigcup_{r=0}^{\ell} S^r)$ we set

$$\begin{aligned}\phi_\tau(s, t; \alpha, \beta; x_1, x, y, z) \\ = \sum_{u \in S} p(\beta, u) st \phi_\tau(s, t; f_1(\alpha, u), f_2(\beta, u); x_1, x, y, z).\end{aligned}$$

Further, we set

$$\begin{aligned}
&\phi_\tau(s, t; f_1(\alpha, u), f_2(\beta, u); x_1, x, y, z) = x_1 \quad \text{if} \quad f_1(\alpha, u) = (b_1, \dots, b_{j*}),\\
&\phi_\tau(s, t; f_1(\alpha, u), f_2(\beta, u); x_1, x, y, z) = x \quad \text{if} \quad f_1(\alpha, u) = A \neq B,\\
&\phi_\tau(s, t; f_1(\alpha, u), f_2(\beta, u); x_1, x, y, z) = y\\
&\qquad\qquad \text{if} \quad f_1(\alpha, u) = B \ \text{ and } \ [B]_{n-m+1}^{n} \neq A,
\end{aligned}$$

and

$$\begin{aligned}
&\phi_\tau(s, t; f_1(\alpha, u), f_2(\beta, u); x_1, x, y, z) = z\\
&\qquad\qquad \text{if} \quad f_1(\alpha, u) = B \ \text{ and } \ [B]_{n-m+1}^{n} = A.
\end{aligned}$$

By solving the above system of equations with respect to

$$\left\{ \phi_\tau(s, t; \alpha, \beta; x_1, x, y, z) \mid (\alpha, \beta) \in CP\left((SPL_0(A) \cup SPL_0(B)) \times \bigcup_{r=0}^{\ell} S^r \right) \right\},$$

we get the generalized p.g.f. of τ with markers (x_1, x, y, z) as the particular solution, $\phi_\tau(s, t; (), (); x_1, x, y, z)$.

Next, we generate the following systems of equations:

$$\begin{cases} \psi_B(t; \alpha, \beta) = \sum_{u \in S} p(\beta, u) t \psi_B(t; f_1(\alpha, u), f_2(\beta, u)) \\ \qquad \text{for every } (\alpha, \beta) \in CP(SPL_0(B) \times S^\ell), \\ \psi_B(t; B, (b_{n-\ell+1}, \dots, b_n)) = 1, \end{cases} \tag{2.1}$$

and

$$\begin{cases} \psi_A(s; \alpha, \beta) = \sum_{u \in S} p(\beta, u) s \psi_A(t; f_1(\alpha, u), f_2(\beta, u)) \\ \qquad \text{for every } (\alpha, \beta) \in CP(SPL_0(A) \times S^\ell), \\ \psi_A(s; A, (a_{m-\ell+1}, \dots, a_m)) = 1. \end{cases} \tag{2.2}$$

By solving (2.1) with respect to $\{\psi_B(t; \alpha, \beta) \mid (\alpha, \beta) \in CP(SPL_0(B) \times S^\ell)\}$, we obtain $\psi_B(t; (b_1, \dots, b_{j*}), (a_{m-\ell+1}, \dots, a_m))$ and $\psi_B(t; A \odot B, (a_{m-\ell+1}, \dots, a_m))$.

Solving (2.2) with respect to $\{\psi_A(s; \alpha, \beta) \mid (\alpha, \beta) \in CP(SPL_0(A) \times S^\ell)\}$, we get $\psi_A(s; B \odot A, (b_{n-\ell+1}, \dots, b_n))$.

Consequently, substituting $x_1 = \psi_B(t; (b_1, \dots, b_{j*}), (a_{m-\ell+1}, \dots, a_m))$, $x = \psi_B(t; A \odot B, (a_{m-\ell+1}, \dots, a_m))$, $y = \psi_A(s; B \odot A, (b_{n-\ell+1}, \dots, b_n))$ and $z = 1$ in $\phi_\tau(s, t; (), (); x_1, x, y, z)$, we obtain $\phi(s, t)$.

3. Illustrative examples

First, we give a simple example of the joint distribution of waiting times for A and B with A completely included in B. Let us consider the waiting times for $A=(1,1)$ and $B=(0,1,1,1)$ in the second order {0,1}-valued Markov chain and let $\tau=\min(W_A,W_B)$. In this case, A comes sooner almost surely. However, we have to check whether the event E_1 occurs or not, when A occurs. Then, in order to calculate the generalized p.g.f. of τ with markers (x_1,x), we write down the system of equations for conditional p.g.f.'s of τ as follows:

$$
\begin{aligned}
\phi_\tau((),()) &= p_{(1)}st\phi_\tau((1),(1)) + p_{(0)}st\phi_\tau((0),(0)),\\
\phi_\tau((1),(1)) &= \frac{p_{(1,1)}}{p_{(1)}}stx + \frac{p_{(1,0)}}{p_{(1)}}st\phi_\tau((0),(1,0)),\\
\phi_\tau((0),(0)) &= \frac{p_{(0,1)}}{p_{(0)}}st\phi_\tau((0,1),(0,1)) + \frac{p_{(0,0)}}{p_{(0)}}st\phi_\tau((0),(0,0)),\\
\phi_\tau((0),(1,0)) &= p_{(1,0),1}st\phi_\tau((0,1),(0,1)) + p_{(1,0),0}st\phi_\tau((0),(0,0)),\\
\phi_\tau((0),(0,0)) &= p_{(0,0),1}st\phi_\tau((0,1),(0,1)) + p_{(0,0),0}st\phi_\tau((0),(0,0)),\\
\phi_\tau((0,1),(0,1)) &= p_{(0,1),1}stx_1 + p_{(0,1),0}st\phi_\tau((0),(1,0)),
\end{aligned}
$$

where $\phi_\tau(\alpha,\beta)$ is short for $\phi_\tau(s,t;\alpha,\beta;x_1,x,y,z)$. By solving the above equations with respect to $\phi_\tau((),())$, $\phi_\tau((1),(1))$, $\phi_\tau((0),(0))$, $\phi_\tau((0),(1,0))$, $\phi_\tau((0,1),(0,1))$ and $\phi_\tau((0),(0,0))$, we obtain

$$
\begin{aligned}
&\phi_\tau(s,t;(),();x_1,x,y,z)\\
&\quad = p_{(1,1)}s^2t^2x\\
&\qquad + \frac{p_{(1,0)}s^2t^2p_{(0,1),1}stx_1\left(p_{(1,0),1}st + \dfrac{p_{(1,0),0}p_{(0,0),1}s^2t^2}{1-p_{(0,0),0}st}\right)}{1-p_{(0,1),0}st\left(p_{(1,0),1}st + \dfrac{p_{(1,0),0}p_{(0,0),1}s^2t^2}{1-p_{(0,0),0}st}\right)}\\
&\qquad + \frac{p_{(0,1)}p_{(0,1),1}s^3t^3x_1}{1-p_{(0,1),0}st\left(p_{(1,0),1}st + \dfrac{p_{(1,0),0}p_{(0,0),1}s^2t^2}{1-p_{(0,0),0}st}\right)}\\
&\qquad + \frac{p_{(0,0)}p_{(0,1),1}s^3t^3x_1\left(\dfrac{p_{(0,0),1}st}{1-p_{(0,0),0}st}\right)}{1-p_{(0,1),0}st\left(p_{(1,0),1}st + \dfrac{p_{(1,0),0}p_{(0,0),1}s^2t^2}{1-p_{(0,0),0}st}\right)}.
\end{aligned}
$$

The system of equations corresponding to (2.1) is

$$
\begin{aligned}
\psi_B((),(1,1)) &= p_{(1,1),1}t\psi_B((),(1,1)) + p_{(1,1),0}t\psi_B((0),(1,0)),\\
\psi_B((0),(1,0)) &= p_{(1,0),1}t\psi_B((0,1),(0,1)) + p_{(1,0),0}t\psi_B((0),(0,0)),\\
\psi_B((0,1),(0,1)) &= p_{(0,1),1}t\psi_B((0,1,1),(1,1)) + p_{(0,1),0}t\psi_B((0),(1,0)),\\
\psi_B((0),(0,0)) &= p_{(0,0),1}t\psi_B((0,1),(0,1)) + p_{(0,0),0}t\psi_B((0),(0,0)),\\
\psi_B((0,1,1),(1,1)) &= p_{(1,1),1}t + p_{(1,1),0}t\psi_B((0),(1,0)).
\end{aligned}
$$

By solving the above equations, we obtain

$$\psi_B((),(1,1)) = \frac{p_{(1,1),0}t}{1-p_{(1,1),1}t}\frac{p_{(0,1),1}p_{(1,1),1}t^2}{C(t)}\left(p_{(1,0),1}t + \frac{p_{(1,0),0}p_{(0,0),1}t^2}{1-p_{(0,0),0}t}\right),$$

and

$$\begin{aligned}\psi_B((0,1,1),(1,1)) &= p_{(1,1),1}t + p_{(1,1),0}t \\ &\qquad \times\left(p_{(1,0),1}t + \frac{p_{(1,0),0}p_{(0,0),1}t^2}{1-p_{(0,0),0}t}\right)\frac{p_{(0,1),1}p_{(1,1),1}t^2}{C(t)},\end{aligned}$$

where

$$C(t) = 1 - (p_{(0,1),1}p_{(1,1),0}t^2 + p_{(0,1),0}t)\left(p_{(1,0),1}t + \frac{p_{(1,0),0}p_{(0,0),1}t^2}{1-p_{(0,0),0}t}\right).$$

Finally, substituting x and x_1 for the above $\psi_B((),(1,1))$ and $\psi_B((0,1,1),(1,1))$, respectively, in $\phi_\tau(s,t;(),();x_1,x,y,z)$, we obtain the joint p.g.f. of (W_A,W_B).

In particular, we further study the joint p.g.f. of W_A and W_B in the sequence of i.i.d. trials for simplicity. By setting $q = 1-p$, $p_{(1,1)} = p^2$, $p_{(1,0)} = pq$, $p_{(0,1)} = pq$, $p_{(0,0)} = q^2$, $p_{(i,j),1} = p$ and $p_{(i,j),0} = q$ for $i,j = 0,1$, we obtain the joint p.g.f. of W_A and W_B as

$$\phi(s,t) = \frac{p^5qs^2t^6}{1-t+p^3qt^4} + \frac{p^2qs^3t^3(pst+1)}{1-qst-pqs^2t^2}\left(pt + \frac{p^3qt^4}{1-qt-pqt^2-p^2qt^3}\right).$$

From the joint p.g.f. we can immediately get the marginal p.g.f.'s

$$\phi(s,1) = \frac{p^2s^2}{1-qs-pqs^2},$$

and

$$\phi(1,t) = \frac{p^3qt^4}{1-t+p^3qt^4}.$$

The next example is the joint probability of waiting times of two patterns $C = (0,1,0,1)$ and $D = (1,0,1,1)$ in $\{0,1\}$-valued independent trials with $P(X_i = 1) = p = 1-q$. The example was introduced by Chen and Zame (1979) in the situation of fair coin tossing games. Assuming $p = \frac{1}{2}$, they showed that $P(W_D < W_C) = \frac{5}{14} < \frac{1}{2}$, whereas $E(W_C) = 20$, $E(W_D) = 18$. From the counterexample, the following intuitive statement was shown to be false, that if $E(W_C) > E(W_D)$ then $P(W_D < W_C) > \frac{1}{2}$. The joint p.g.f. of W_C and W_D can be written as

$$\begin{aligned}\phi(s,t) = p^2qs^4t^4 &\Big\{(-p^4q^4s^7t^3 - p^3q^3s^5t + p^3q^2s^4 + p^4q^3s^6t^2)\frac{1}{\xi(s)} \\ &\qquad + (pqt + p^3q^2s^3t^4)\frac{\zeta(t)}{\eta(t)}\Big\} \\ &\times \{1 - st + pqs^2t^2 - pq^2s^3t^3 - p^3q^3s^6t^6\}^{-1},\end{aligned}$$

where

$$\xi(s) = 1 - s + pqs^2 - pqs^3 + p^2q^2s^4,$$
$$\eta(t) = -1 + t - p^2qt^3 + p^2q^2t^4,$$
$$\zeta(t) = -1 + t - pqt^2 + pq^2t^3.$$

The joint probability function of (W_C, W_D) is shown in Fig. 1. From the joint

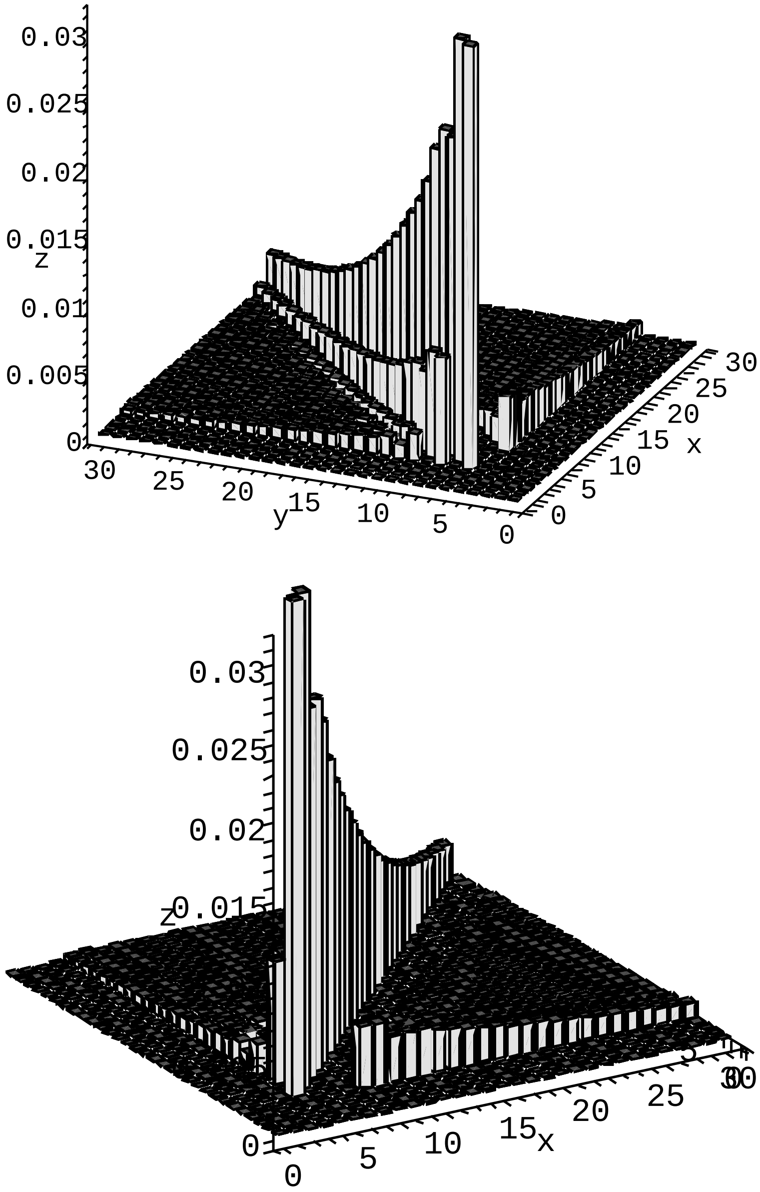

Figure 1. The joint probability of waiting times of two patterns $C = (0,1,0,1)$ and $D = (1,0,1,1)$ in {0,1}-valued independent trials with $P(X_i = 1) = \frac{1}{2}$. These two graphs are the same ones from different viewpoints.

p.g.f., we can derive exact means and variances of W_C and W_D and the covariance and correlation coefficient of W_C and W_D. Since they are too long to be written here, we give only the values of them for $p = \frac{1}{2}$; $E(W_C) = 20$, $E(W_D) = 18$, $V(W_C) = 276$, $V(W_D) = 210$, $\mathrm{Cov}(W_C, W_D) = \frac{384}{7}$ and $\rho = \frac{32\sqrt{1610}}{5635}$.

Acknowledgements

The authors remember that Professor Akaike usually insisted on the importance of statistical modeling in the period of early eighties. His great insight had a good impact on the authors' research work toward exact discrete distribution theory, which has been developed considerably by many researchers since then and nowadays enables us to treat practical modeling. The authors would like to thank Professor Akaike at this opportunity. Their thanks also go to Professor Taniguchi and the editorial committee for the invitation to the special volume. The research was partially supported by the ISM Cooperative Research Program (2007-ISM-CRP-2044) and by the Grant-in-Aid for Scientific Research (C) of the JSPS (Grant Number 16500183 and Grant Number 18540148).

References

Balakrishnan, N. and Koutras, M. V. (2002). *Runs and Scans with Applications*, Wiley, New York.

Chen, R. and Zame, A. (1979). On fair coin-tossing games, *Journal of Multivariate Analysis*, **9**, 150–156.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner and later problems for Bernoulli trials: Frequency and run quotas, *Statistics & Probability Letters*, **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 Chang, Y. M. (2002). On probability generating functions for waiting time distributions of compound patterns in a sequence of multistate trials, *Journal of Applied Probability*, **39**, 70–80.

Fu, J. C. and Lou, W. Y. W. (2003). *Distribution Theory of Runs and Patterns and its Applications: A Finite Markov Chain Imbedding Approach*, World Scientific Publishing Co.