AISM (2008) 60:173–191
DOI 10.1007/s10463-006-0080-6

Bivariate Markov chain embeddable variables of polynomial type

M. V. Koutras · S. Bersimis · D. L. Antzoulakos

Received: 18 October 2005 / Revised: 10 February 2006 / Published online: 18 October 2006

Abstract The primary aim of the present article is to provide a general framework for investigating the joint distribution of run length accumulating/enumerating variables by the aid of a Markov chain embedding technique. To achieve that we introduce first a class of bivariate discrete random variables whose joint distribution can be described by the aid of a Markov chain and develop formulae for their joint probability mass function, generating functions and moments. The results are then exploited for the derivation of the distribution of a bivariate run-related statistic. Finally, some interesting uses of our results in reliability theory and educational psychology are highlighted.

Keywords Success runs · Run lengths · Markov chains · Consecutive-*r*-out-of-*n*: *F* system

1 Introduction

In the past years, special research interest has been drawn on problems associated with the frequency of occurrences of runs of like symbols in a sequence of binary or multistate trials. The more general problem of studying the number of occurrences of arbitrary patterns has also attracted considerable research interest. For the evaluation of the stochastic behaviour of run (or pattern) enumerating random variables, several approaches have been suggested based

Research supported by General Secretary of Research and Technology of Greece under grand PENED 2001.

M. V. Koutras (✉) · S. Bersimis · D. L. Antzoulakos
Department of Statistics and Insurance Science, University of Piraeus, Karaoli & Dimitriou 80, Piraeus 18534, Greece
e-mail: mkoutras@unipi.gr

on direct combinatorial considerations, on renewal theory and recurrent events, on martingale techniques, on generating function theory etc (see e.g., Feller, 1968, Gani, 2003, Guibas and Odlyzko, 1981, Leslie, 1967, Mood, 1940, Rajarshi, 1974).

Due to the wide variety of run (pattern) related statistics, it would be helpful if one could fall back on some unifying principle in their study. The use of Markov chain techniques has highly facilitated this task. Recent work in this area is outlined in the monographs by Balakrishnan and Koutras (2002), Fu and Lou (2003), and Glaz et al. (2001).

While a lot of publications have dealt with statistics related to the number of runs, only a few of them take into account the exact run length. Agin and Godbole (1992) suggested using run lengths of variable size for establishing non parametric randomness tests. An approach of similar flavour was considered earlier by O'Brien and Dyck (1985). Motivated by those works, Antzoulakos et al. (2003) studied the distribution of a statistic accumulating the run lengths of "reasonably long" strands of like elements (successes) in a sequence of binary trials. To achieve that, they exploited a proper Markov chain embedding technique, thereby creating a powerful working environment for investigating a much wider range of run/pattern related statistics. See also Lou (2003) who also used a Markov chain embedding technique to investigate the same statistic motivated by a molecular biology application.

An interesting extension of the aforementioned models would arise by looking at multivariate models related to variables accumulating several types of run lengths (e.g. lengths of failure runs and lengths of success runs). Such generalizations besides their theoretical appeal, will also provide appropriate probability models for tackling a variety of attractive applications. In order to have a wide framework that could accommodate several accumulating/enumerating schemes of run lengths, we shall first build an appropriate Markov chain approach and subsequently exploit it to study a specific multivariate run accumulating statistic.

In Sect. 2 we introduce the concept of a bivariate Markov chain embeddable variable of polynomial type (BMVP). In the same section, several compact and computationally tractable formulae are deduced for the evaluation of the joint probability mass function, generating functions and moments of a BMVP. A brief discussion is also included, detailing the possible extension of the bivariate model to $s \geq 2$ dimensions. In Sect. 3, we illustrate how the general results obtained in Sect. 2 can be applied in the study of the joint distributions of run-accumulating variables. Finally, in Sect. 4, we highlight the use of these distributions in two applied research areas: reliability theory and educational psychology.

2 General results

Fu and Koutras (1994) developed a unified method for capturing the exact distribution of the number of runs of specified length by employing a Markov chain embedding technique. Koutras and Alexandrou (1995) refined the method and expressed these distributions in terms of multidimensional binomial type probability vectors by introducing the concept of Markov chain embeddable variables of binomial type. Fu (1996) extended the original method to cover the case of arbitrary patterns (instead of runs) whereas Koutras (1997) treated several waiting time problems within this framework. Finally Doi and Yamamoto (1998) and Han and Aki (1999) considered the case of multivariate run related distributions and offered simple solutions by exploiting proper extensions of the Markov chain embedding technique (for an illustrative presentation of this method see Koutras, 2003).

Recently, Antzoulakos et al. (2003) generalized the method of Koutras and Alexandrou (1995) by introducing the concept of Markov chain embeddable variables of polynomial type (MVP). In the sequel, we extend this method in two dimensions, thereby obtaining multivariate analogues of the MVP's.

Definition 1 *A bivariate discrete random variable* $(X_n^{(1)}, X_n^{(2)})$ *will be called a* BMVP *if*

(a) there exists a Markov chain $\{Y_t, t \geq 0\}$ defined on a state space Ω which can be partitioned as

$$\Omega = \bigcup_{x_1,x_2 \geq 0} C_{x_1,x_2}, \quad C_{x_1,x_2} = \{c_{x_1,x_2;0}, c_{x_1,x_2;1}, \dots, c_{x_1,x_2;s-1}\}$$

(b) there exist two positive integers m_1 and m_2 such that

$$\Pr\left(Y_t \in C_{y_1,y_2} \mid Y_{t-1} \in C_{x_1,x_2}\right) = 0, \quad t \geq 1,$$

for all $(y_1, y_2) \notin \{(x_1 + u, x_2) : 0 \leq u \leq m_1\} \cup \{(x_1, x_2 + v) : 0 \leq v \leq m_2\}$

(c) the joint probability mass function of $(X_n^{(1)}, X_n^{(2)})$ can be captured by considering the projection of the probability space of Y_n onto C_{x_1,x_2} i.e.

$$\Pr(X_n^{(1)} = x_1, \ X_n^{(2)} = x_2) = \Pr(Y_n \in C_{x_1,x_2}), \quad n \geq 0, \ x_1, x_2 \geq 0.$$

For $m_1 = m_2 = 1$, Definition 1 reduces to the bivariate version of the Markov chain embeddable vector of multinomial type introduced by Han and Aki (1999).

Roughly speaking, a BMVP is characterized by the following property: once the chain enters C_{x_1,x_2}, the feasible one step transitions lead either to the same subclass C_{x_1,x_2}, or to one of the subclasses $C_{x_1+1,x_2}, C_{x_1+2,x_2}, \dots, C_{x_1+m_1,x_2}$, or to one of the subclasses $C_{x_1,x_2+1}, C_{x_1,x_2+2}, \dots, C_{x_1,x_2+m_2}$.

The distribution of a BMVP can be easily captured if we have at hand the following quantities

- the initial probabilities

$$\boldsymbol{\pi}_{x_1,x_2} = \left(\Pr(Y_0 = c_{x_1,x_2;0}), \Pr(Y_0 = c_{x_1,x_2;1}), \dots, \Pr(Y_0 = c_{x_1,x_2;s-1})\right)$$

- the within states one step transition probability matrix

$$A_{t,0}(x_1,x_2) = \left(\Pr\left(Y_t = c_{x_1,x_2;j,j'} \mid Y_{t-1} = c_{x_1,x_2;j}\right)\right)_{s\times s}$$

- the between states one step transition probability matrices

$$A^{(1)}_{t,u}(x_1,x_2) = \left(\Pr\left(Y_t = c_{x_1+u,x_2;j,j'} \mid Y_{t-1} = c_{x_1,x_2;j,j'}\right)\right)_{s\times s}, \quad 1 \le u \le m_1$$
$$A^{(2)}_{t,v}(x_1,x_2) = \left(\Pr\left(Y_t = c_{x_1,x_2+v;j,j'} \mid Y_{t-1} = c_{x_1,x_2;j,j'}\right)\right)_{s\times s}, \quad 1 \le v \le m_2$$

$(t \ge 1,\ x_1,x_2 \ge 0,\ 0 \le j,j' \le s-1\ )$.

It is clear that, for all $t \ge 1$, the matrix

$$A_{t,0}(x_1,x_2) + \sum_{u=1}^{m_1} A^{(1)}_{t,u}(x_1,x_2) + \sum_{v=1}^{m_2} A^{(2)}_{t,v}(x_1,x_2)$$

is stochastic.

On introducing the probability (row) vectors

$$\mathbf{f}_t(x_1,x_2) = \left(\Pr(Y_t = c_{x_1,x_2;0}), \Pr(Y_t = c_{x_1,x_2;1}), \ldots, \Pr(Y_t = c_{x_1,x_2;s-1})\right),$$

for $t \ge 0$ and $x_1,x_2 \ge 0$, it follows directly from condition (c) of Definition 1 that the joint probability mass function of $(X^{(1)}_n, X^{(2)}_n)$ is given by

$$\Pr(X^{(1)}_n = x_1,\ X^{(2)}_n = x_2) = \mathbf{f}_n(x_1,x_2)\mathbf{1}', \quad n \ge 0,\ \ x_1,x_2 \ge 0,$$

where $\mathbf{1} = (1,1,\ldots,1)$ denotes the row vector of $\mathbb{R}^s$ with all its entries being 1. Finally, the convention $\Pr(X^{(1)}_0 = 0, X^{(2)}_0 = 0) = 1$ implies that

$$\boldsymbol{\pi}_{0,0}\mathbf{1}' = \mathbf{f}_0(0,0)\mathbf{1}' = \left(\Pr(Y_0 = c_{0,0;0}), \Pr(Y_0 = c_{0,0;1}), \ldots, \Pr(Y_0 = c_{0,0;s-1})\right)\mathbf{1}' = 1$$

and

$$\boldsymbol{\pi}_{x_1,x_2}\mathbf{1}' = \mathbf{f}_0(x_1,x_2)\mathbf{1}' = 0, \quad \text{if } x_1 \ge 1 \quad \text{or } x_2 \ge 1.$$

In the theorems that follow, we provide several results that facilitate the evaluation of the joint distribution of a BMVP. The first one offers an efficient recursive scheme for the computation of the vector $\mathbf{f}_t(x_1,x_2)$. The next two theorems present formulae for the joint probability generating function while the last one deals with the moments of a BMVP.

Theorem 1 *The sequence of vectors* $\mathbf{f}_t(x_1,x_2)$ *satisfies the recurrence relation*

$$\mathbf{f}_t(x_1,x_2) = \mathbf{f}_{t-1}(x_1,x_2)A_{t,0}(x_1,x_2) + \sum_{u=1}^{\min(x_1,\,m_1)} \mathbf{f}_{t-1}(x_1-u,x_2)A_{t,u}^{(1)}(x_1-u,x_2)$$
$$+ \sum_{v=1}^{\min(x_2,\,m_2)} \mathbf{f}_{t-1}(x_1,x_2-v)A_{t,v}^{(2)}(x_1,x_2-v), \quad t\geq 1, \ \ x_1,x_2\geq 0.$$

Proof Let $t \geq 1$, $x_1,x_2 \geq 0$ and $0 \leq j \leq s-1$. In view of the total probability theorem we may write

$$\begin{aligned}
&\Pr(Y_t = c_{x_1,x_2;j}) \\
&\quad = \sum_{r=0}^{s-1} \Pr\left(Y_t = c_{x_1,x_2;j} \mid Y_{t-1} = c_{x_1,x_2;r}\right) \Pr\left(Y_{t-1} = c_{x_1,x_2;r}\right) \\
&\qquad + \sum_{u=1}^{\min(x_1,\,m_1)} \sum_{r=0}^{s-1} \Pr\left(Y_t = c_{x_1,x_2;j} \mid Y_{t-1} = c_{x_1-u,x_2;r}\right)\Pr\left(Y_{t-1} = c_{x_1-u,x_2;r}\right) \\
&\qquad + \sum_{v=1}^{\min(x_2,\,m_2)} \sum_{r=0}^{s-1} \Pr\left(Y_t = c_{x_1,x_2;j} \mid Y_{t-1} = c_{x_1,x_2-v;r}\right)\Pr\left(Y_{t-1} = c_{x_1,x_2-v;r}\right).
\end{aligned}$$

Expressing the conditional probabilities in terms of within and between states transition probability matrices we deduce

$$\begin{aligned}
&\Pr(Y_t = c_{x_1,x_2;j,j'}) \\
&\quad = \sum_{r=0}^{s-1} \mathbf{e}_{r+1}A_{t,0}(x_1,x_2)\mathbf{e}'_{j+1} \Pr\left(Y_{t-1} = c_{x_1,x_2;r}\right) \\
&\qquad + \sum_{u=1}^{\min(x_1,\,m_1)} \sum_{r=0}^{s-1} \mathbf{e}_{r+1}A_{t,u}^{(1)}(x_1-u,x_2)\mathbf{e}'_{j+1} \Pr\left(Y_{t-1} = c_{x_1-u,x_2;r}\right) \\
&\qquad + \sum_{v=1}^{\min(x_2,\,m_2)} \sum_{r=0}^{s-1} \mathbf{e}_{r+1}A_{t,v}^{(2)}(x_1,x_2-v)\mathbf{e}'_{j+1} \Pr\left(Y_{t-1} = c_{x_1,x_2-v;r}\right) \\
&\quad = \mathbf{f}_{t-1}(x_1,x_2)A_{t,0}(x_1,x_2)\mathbf{e}'_{j+1} \\
&\qquad + \sum_{u=1}^{\min(x_1,\,m_1)} \mathbf{f}_{t-1}(x_1-u,x_2)A_{t,u}^{(1)}(x_1-u,x_2)\mathbf{e}'_{j+1} \\
&\qquad + \sum_{v=1}^{\min(x_2,\,m_2)} \mathbf{f}_{t-1}(x_1,x_2-v)A_{t,v}^{(2)}(x_1,x_2-v)\mathbf{e}'_{j+1}
\end{aligned}$$

($\mathbf{e}_i$ denote the unit row vectors of $\mathbb{R}^s$), and the proof is complete. □

Let us now denote by $\varphi_t(z_1, z_2)$ and $\Phi(z_1, z_2; w)$ the single and double generating functions of the bivariate random variable $(X_n^{(1)}, X_n^{(2)})$ respectively, that is

$$\varphi_t(z_1, z_2) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \Pr(X_t^{(1)} = x_1,\ X_t^{(2)} = x_2) z_1^{x_1} z_2^{x_2},$$

$$\Phi(z_1, z_2; w) = \sum_{t=0}^{\infty} \varphi_t(z_1, z_2) w^t.$$

In addition, let $\boldsymbol{\varphi}_t(z_1, z_2)$ and $\boldsymbol{\Phi}(z_1, z_2; w)$ stand for the single (row) and double (row) vector generating functions of $\mathbf{f}_t(x_1, x_2)$, respectively, that is

$$\boldsymbol{\varphi}_t(z_1, z_2) = \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \mathbf{f}_t(x_1, x_2) z_1^{x_1} z_2^{x_2}, \quad \boldsymbol{\Phi}(z_1, z_2; w) = \sum_{t=0}^{\infty} \boldsymbol{\varphi}_t(z_1, z_2) w^t.$$

It is then clear that

$$\varphi_t(z_1, z_2) = \boldsymbol{\varphi}_t(z_1, z_2)\mathbf{1}', \quad \Phi(z_1, z_2; w) = \boldsymbol{\Phi}(z_1, z_2; w)\mathbf{1}'$$

for all $t \geq 0$. Moreover, the obvious identity $\boldsymbol{\varphi}_0(z_1, z_2) = \boldsymbol{\pi}_{0,0}$ implies that $\varphi_0(z_1, z_2) = \boldsymbol{\varphi}_0(z_1, z_2)\mathbf{1}' = 1$, a result that is in compliance with the convention $\Pr(X_0^{(1)} = 0, X_0^{(2)} = 0) = 1$.

We mention that, it is the rule rather than the exception that the transition probability matrices associated to a BMVP do not depend on (x_1, x_2), that is $A_{t,0}(x_1, x_2) = A_{t,0}$, $A_{t,u}^{(1)}(x_1, x_2) = A_{t,u}^{(1)}$ $(1 \leq u \leq m_1)$, $A_{t,v}^{(2)}(x_1, x_2) = A_{t,v}^{(2)}$ $(1 \leq v \leq m_2)$ for all $t \geq 1$ and $x_1, x_2 \geq 0$. In this case, the vector generating function $\boldsymbol{\varphi}_t(z_1, z_2)$ can be expressed in the form of a matrix product as the following theorem indicates.

Theorem 2 *If* $A_{t,0}(x_1, x_2) = A_{t,0}$, $A_{t,u}^{(1)}(x_1, x_2) = A_{t,u}^{(1)}$, $1 \leq u \leq m_1$, *and* $A_{t,v}^{(2)}(x_1, x_2) = A_{t,v}^{(2)}$, $1 \leq v \leq m_2$, *for all* $t \geq 1$ *and* $x_1, x_2 \geq 0$, *then the (single) vector generating function of* $\mathbf{f}_t(x_1, x_2)$ *can be expressed in the form*

$$\boldsymbol{\varphi}_t(z_1, z_2) = \boldsymbol{\pi}_{0,0} \prod_{r=1}^{t} \left(A_{r,0} + \sum_{u=1}^{m_1} A_{r,u}^{(1)} z_1^u + \sum_{v=1}^{m_2} A_{r,v}^{(2)} z_2^v \right), \quad t \geq 1.$$

Proof Let us first decompose the (single) vector generating function of $\mathbf{f}_t(x_1, x_2)$ as follows

$$\boldsymbol{\varphi}_t(z_1, z_2) = \mathbf{f}_t(0,0) + \sum_{i=1}^{6} \sum_{S_i} \mathbf{f}_t(x_1, x_2) z_1^{x_1} z_2^{x_2}$$

where

$$
\begin{aligned}
S_1 &= \{(x_1,x_2) : x_1 = 0,\ x_2 \geq 1\},\\
S_2 &= \{(x_1,x_2) : x_1 \geq 1,\ x_2 = 0\},\\
S_3 &= \{(x_1,x_2) : 1 \leq x_1 \leq m_1,\ 1 \leq x_2 \leq m_2\},\\
S_4 &= \{(x_1,x_2) : 1 \leq x_1 \leq m_1,\ x_2 \geq m_2+1\},\\
S_5 &= \{(x_1,x_2) : x_1 \geq m_1+1,\ 1 \leq x_2 \leq m_2\},\\
S_6 &= \{(x_1,x_2) : x_1 \geq m_1+1,\ x_2 \geq m_2+1\}.
\end{aligned}
$$

Exploiting the recurrences stated in Theorem 1 we may rewrite the sum $\sum_{S_4} \mathbf{f}_t(x_1,x_2) z_1^{x_1} z_2^{x_2}$ in the form

$$
\begin{aligned}
&\sum_{x_1=1}^{m_1} \sum_{x_2=m_2+1}^{\infty} \mathbf{f}_t(x_1,x_2) z_1^{x_1} z_2^{x_2}\\
&\quad= \sum_{x_1=1}^{m_1} \sum_{x_2=m_2+1}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,0} z_1^{x_1} z_2^{x_2}\\
&\qquad+ \sum_{x_1=1}^{m_1} \sum_{x_2=m_2+1}^{\infty} \sum_{u=1}^{x_1} \mathbf{f}_{t-1}(x_1-u,x_2) A_{t,u}^{(1)} z_1^{x_1} z_2^{x_2}\\
&\qquad+ \sum_{x_1=1}^{m_1} \sum_{x_2=m_2+1}^{\infty} \sum_{v=1}^{m_2} \mathbf{f}_{t-1}(x_1,x_2-v) A_{t,v}^{(2)} z_1^{x_1} z_2^{x_2}\\
&\quad= \sum_{x_1=1}^{m_1} \sum_{x_2=m_2+1}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,0} z_1^{x_1} z_2^{x_2}\\
&\qquad+ \sum_{u=1}^{m_1} z_1^{u} \left(\sum_{x_1=0}^{m_1-u} \sum_{x_2=m_2+1}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,u}^{(1)} z_1^{x_1} z_2^{x_2} \right)\\
&\qquad+ \sum_{v=1}^{m_2} z_2^{v} \left(\sum_{x_1=1}^{m_1} \sum_{x_2=m_2-v+1}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,v}^{(2)} z_1^{x_1} z_2^{x_2} \right).
\end{aligned}
$$

Applying a similar technique for the remaining summands (replace also $\mathbf{f}_t(0,0)$ by $\mathbf{f}_{t-1}(0,0)A_{t,0}$ when necessary), we may easily deduce the expression

$$
\begin{aligned}
\boldsymbol{\varphi}_t(z_1,z_2) &= \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,0} z_1^{x_1} z_2^{x_2}\\
&\quad+ \sum_{u=1}^{m_1} z_1^{u} \left(\sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,u}^{(1)} z_1^{x_1} z_2^{x_2} \right)
\end{aligned}
$$

$$+\sum_{v=1}^{m_2} z_2^v \left(\sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \mathbf{f}_{t-1}(x_1,x_2) A_{t,v}^{(2)} z_1^{x_1} z_2^{x_2} \right)$$

which leads to the recursive formula

$$\boldsymbol{\varphi}_t(z_1,z_2) = \boldsymbol{\varphi}_{t-1}(z_1,z_2)\left(A_{t,0} + \sum_{u=1}^{m_1} A_{t,u}^{(1)} z_1^u + \sum_{v=1}^{m_2} A_{t,v}^{(2)} z_2^v \right), \quad t \geq 1.$$

A repeated application of the last formula, yields the desired result. □

In the case of a homogeneous BMVP, i.e. when $A_{t,u}^{(1)}(x_1,x_2)$, $A_{t,v}^{(2)}(x_1,x_2)$ do not depend on t, x_1, x_2 we have the next theorem.

Theorem 3 *If* $A_{t,0}(x_1,x_2) = A_0$, $A_{t,u}^{(1)}(x_1,x_2) = A_u^{(1)}$, $1 \leq u \leq m_1$, *and* $A_{t,v}^{(2)}(x_1,x_2) = A_v^{(2)}$, $1 \leq v \leq m_2$, *for all* $t \geq 1$ *and* x_1, $x_2 \geq 0$, *then the double vector generating function of* $\mathbf{f}_t(x_1,x_2)$ *can be expressed as*

$$\boldsymbol{\Phi}(z_1,z_2;w) = \boldsymbol{\pi}_{0,0}\left[I - w\left(A_0 + \sum_{u=1}^{m_1} A_u^{(1)} z_1^u + \sum_{v=1}^{m_2} A_v^{(2)} z_2^v \right)\right]^{-1}$$

where I *is the identity* $s \times s$ *matrix.*

Proof Follows readily from Theorem 2 on observing that

$$\begin{aligned}\boldsymbol{\Phi}(z_1,z_2;w) &= \sum_{t=0}^{\infty} \boldsymbol{\varphi}_t(z_1,z_2) w^t \\ &= \boldsymbol{\pi}_{0,0} \sum_{t=0}^{\infty} \left(w\left(A_0 + \sum_{u=1}^{m_1} A_u^{(1)} z_1^u + \sum_{v=1}^{m_2} A_v^{(2)} z_2^v \right)\right)^t \\ &= \boldsymbol{\pi}_{0,0}\left[I - w\left(A_0 + \sum_{u=1}^{m_1} A_u^{(1)} z_1^u + \sum_{v=1}^{m_2} A_v^{(2)} z_2^v \right)\right]^{-1}\end{aligned}$$

the last equality being valid in an appropriate neighbourhood of zero for w. □

For a homogeneous BMVP, the next theorem provides compact formulae for the evaluation of the means $E[X_t^{(1)}], E[X_t^{(2)}]$, $E[X_t^{(1)} X_t^{(2)}]$ and their corresponding generating functions.

Theorem 4 *If* $A_{t,0}(x_1,x_2)=A_0$, $A^{(1)}_{t,u}(x_1,x_2)=A^{(1)}_u, 1\le u\le m_1$, *and* $A^{(2)}_{t,v}(x_1,x_2)=A^{(2)}_v, 1\le v\le m_2$, *for all* $t\ge 1$ *and* $x_1,x_2\ge 0$, *then*

$$E[X_t^{(j)}]=\boldsymbol{\pi}_{0,0}\sum_{r=1}^{t}B^{r-1}D_j\mathbf{1}',\quad j=1,2$$

$$E[X_t^{(1)}X_t^{(2)}]=\boldsymbol{\pi}_{0,0}\sum_{r=1}^{t}\left(\sum_{i=1}^{r-1}B^{i-1}D_2B^{r-1-i}D_1+B^{r-1}D_1\sum_{i=1}^{t-r}B^{i-1}D_2\right)\mathbf{1}'$$

$$M_j(w)=\sum_{t=1}^{\infty}E[X_t^{(j)}]w^t=\frac{w}{1-w}\boldsymbol{\pi}_{0,0}(I-wB)^{-1}D_j\mathbf{1}',\quad j=1,2$$

$$\begin{aligned}M_{1,2}(w)&=\sum_{t=1}^{\infty}E[X_t^{(1)}X_t^{(2)}]w^t\\&=\frac{w^2}{1-w}\boldsymbol{\pi}_{0,0}(I-wB)^{-1}[D_1(I-wB)^{-1}D_2+D_2(I-wB)^{-1}D_1]\mathbf{1}'\end{aligned}$$

where

$$B=A_0+\sum_{u=1}^{m_1}A_u^{(1)}+\sum_{v=1}^{m_2}A_v^{(2)},\quad D_1=\sum_{u=1}^{m_1}uA_u^{(1)},\quad D_2=\sum_{v=1}^{m_2}vA_v^{(2)}.$$

Proof Observe first that, for arbitrary square matrices Q_i, $0\le i\le k$, the following identity holds true

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\sum_{i=0}^{k}z^iQ_i\right)^t=\sum_{r=1}^{t}\left[\left(\sum_{i=0}^{k}z^iQ_i\right)^{r-1}\left(\sum_{i=1}^{k}iQ_iz^{i-1}\right)\left(\sum_{i=0}^{k}z^iQ_i\right)^{t-r}\right].$$

Since we are dealing with a homogeneous BMVP, $\boldsymbol{\varphi}_t(z_1,z_2)$ takes on the form

$$\boldsymbol{\varphi}_t(z_1,z_2)=\boldsymbol{\pi}_{0,0}\left(A_0+\sum_{u=1}^{m_1}A_u^{(1)}z_1^u+\sum_{v=1}^{m_2}A_v^{(2)}z_2^v\right)^t,\quad t\ge 1 \tag{1}$$

and making use of the last identity we obtain

$$E[X_t^{(1)}]=\frac{\partial}{\partial z_1}\left[\boldsymbol{\varphi}_t(z_1,z_2)\mathbf{1}'\right]\Big|_{z_1=z_2=1}=\boldsymbol{\pi}_{0,0}\sum_{r=1}^{t}B^{r-1}D_1B^{t-r}\mathbf{1}'$$

where

$$B=A_0+\sum_{u=1}^{m_1}A_u^{(1)}+\sum_{v=1}^{m_2}A_v^{(2)},\quad D_1=\sum_{u=1}^{m_1}uA_u^{(1)}.$$

The desired expression for $E[X_t^{(1)}]$ is readily ascertained by recalling that matrix B is stochastic and therefore $B^i\mathbf{1}' = \mathbf{1}'$, for $i = 0, 1, 2, \ldots, t-1$.

The generating function $M_1(w)$ of the means $E[X_t^{(1)}], t \geq 1$, may be written as

$$M_1(w) = \boldsymbol{\pi}_{0,0} \sum_{t=1}^{\infty} \sum_{r=1}^{t} B^{r-1} D_1 w^t \mathbf{1}' = \boldsymbol{\pi}_{0,0} \sum_{r=1}^{\infty} (wB)^{r-1} \sum_{t=r}^{\infty} w^{t-r+1} D_1 \mathbf{1}'$$

and the final formula is effortlessly established by virtue of

$$\sum_{r=1}^{\infty} (wB)^{r-1} = (I - wB)^{-1}.$$

Proceeding in a similar way we may readily obtain analogous expressions for $E[X_t^{(2)}]$ and $M_2(w)$.

The evaluation of the mean $E[X_t^{(1)} X_t^{(2)}]$ may be accomplished by the aid of the formula

$$E[X_t^{(1)} X_t^{(2)}] = \frac{\partial^2}{\partial z_1 \partial z_2} \left[\boldsymbol{\varphi}_t(z_1, z_2) \mathbf{1}' \right] \Big|_{z_1 = z_2 = 1}.$$

Differentiating (1) with respect to z_1, z_2 we obtain

$$E[X_t^{(1)} X_t^{(2)}] = \boldsymbol{\pi}_{0,0} \sum_{r=1}^{t} \left(\sum_{i=1}^{r-1} B^{i-1} D_2 B^{r-1-i} D_1 B^{t-r} + B^{r-1} D_1 \sum_{i=1}^{t-r} B^{i-1} D_2 B^{t-r-i} \right) \mathbf{1}'$$

and the desired formula for $E[X_t^{(1)} X_t^{(2)}]$ follows immediately by taking into account that matrix B is stochastic.

Finally, the generating function $M_{1,2}(w)$ of the means $E[X_t^{(1)} X_t^{(2)}], t \geq 1$, can be easily deduced by exploiting the next two identities

$$\sum_{t=1}^{\infty} \sum_{r=1}^{t} \sum_{i=1}^{r-1} B^{i-1} D_2 B^{r-1-i} w^t = \frac{w^2}{1-w} (I - wB)^{-1} D_2 (I - wB)^{-1},$$

$$\sum_{t=1}^{\infty} \sum_{r=1}^{t} \sum_{i=1}^{t-r} B^{r-1} D_1 B^{i-1} w^t = \frac{w^2}{1-w} (I - wB)^{-1} D_1 (I - wB)^{-1}$$

which are readily ascertainable after some elementary algebra. □

In closing we mention that all the abovementioned definitions and results can be effortlessly generalized to the multivariate case. The details are left to the interested reader.

3 A special case

Consider a sequence of Bernoulli trials $Z_1, Z_2, \ldots$ with success probabilities $p_t = P(Z_t = 1)$, and failure probabilities $q_t = P(Z_t = 0) = 1 - p_t$, $t \geq 1$, and let n, k and r be any positive integers with $n \geq \max(k, r)$. Denote by $X_{n,k}^{(1)}$ the sum of run lengths of the success runs of length at least k observed in the sequence $Z_1, Z_2, \ldots, Z_n$ and by $X_{n,r}^{(2)}$ the number of non-overlapping failure runs of length r in the same sequence.

The random variable $X_{n,r}^{(2)}$ is a Markov chain embeddable variable of binomial type and has been thoroughly studied within this framework by Koutras and Alexandrou (1995) while $X_{n,r}^{(1)}$ is a (univariate) MVP (see, Antzoulakos et al., 2003). However, the joint distribution of $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$ has not, to the best of our knowledge, been discussed hitherto in the literature. The general results derived in the previous section, offer a quite effective framework to evaluate the joint probability mass function of $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$ and the respective generating functions.

In order to view the bivariate random variable $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$ as a *BMVP*, we use the state space $\Omega = \bigcup_{x_1, x_2 \geq 0} C_{x_1, x_2}$ where

$$C_{x_1,x_2} = \{(x_1, x_2; i, 1) : 1 \leq i \leq k\} \cup \{(x_1, x_2; j, 0) : 0 \leq j \leq r - 1\}$$

for all $x_1, x_2 \geq 0$. A state of the form $(x_1, x_2; \cdot, \cdot)$ indicates that, in the first outcomes of the sequence under investigation (a) the observed sum of the exact lengths of runs of k or more consecutive successes is x_1, and (b) the number of non-overlapping failure runs of length r is x_2.

The states of the form $(x_1, x_2; i, 1)$ are used for keeping track of success runs while the states $(x_1, x_2; j, 0)$ for keeping track of failure runs. More specifically, we define the Markov chain $\{Y_t, t \geq 0\}$ as follows:

- If the first t outcomes of the binary sequence under investigation are of the form $1001 \cdots 0\underbrace{11 \cdots 1}_{c \geq 1}$, we set $Y_t = (x_1, x_2; i, 1)$, $1 \leq i \leq k$, where

$$i = \begin{cases} c, \text{ if } c = 1, \ldots, k - 1 \\ k, \text{ if } c \geq k. \end{cases}$$

- If the first t outcomes of the binary sequence under investigation are of the form $1001 \cdots 1\underbrace{00 \cdots 0}_{c \geq 1}$, we set $Y_t = (x_1, x_2; j, 0)$, $0 \leq j \leq r–1$, where $j = c (\mathrm{mod}\, r)$.

It is apparent that, once the chain enters C_{x_1,x_2}, the one step transitions may lead only to the subclasses C_{x_1,x_2}, C_{x_1+1,x_2}, C_{x_1+k,x_2} or C_{x_1,x_2+1}. Hence the random variable $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$ belongs to the class of BMVP ($m_1 = k$, $m_2 = 1$). The transition probability matrices $A_{t,0}$, $A_{t,i}^{(1)}$ $(i = 1, 2, \ldots, k)$ and $A_{t,1}^{(2)}$ can be easily identified by observing that

(a) if $Y_t = (x_1, x_2; k, 1)$ the feasible one step transitions of the chain lead either to substate $(x_1 + 1, x_2; k, 1)$ (if $Z_{t+1} = 1$) or to substate $(x_1, x_2; 1, 0)$ (if $Z_{t+1} = 0$), and
(b) if $Y_t = (x_1, x_2; r - 1, 0)$ the feasible one step transitions of the chain lead either to substate $(x_1, x_2 + 1; 0, 0)$ (if $Z_{t+1} = 0$) or to substate $(x_1, x_2; 1, 0)$ (if $Z_{t+1} = 1$).

Therefore, $A_{t,0}$ will be given by

$$A_{t,0} = \left[\begin{array}{ccccccc|cccccc}
0 & p_t & 0 & \cdots & 0 & 0 & 0 & q_t & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & p_t & \cdots & 0 & 0 & 0 & q_t & 0 & 0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & p_t & 0 & 0 & q_t & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & q_t & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & 0 & q_t & 0 & 0 & \cdots & 0 & 0 \\
\hline
0 & p_t & 0 & \cdots & 0 & 0 & 0 & 0 & q_t & 0 & \cdots & 0 & 0 \\
0 & p_t & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & q_t & & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & p_t & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & q_t \\
0 & p_t & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0
\end{array}\right]_{(k+r)\times(k+r)}$$

with the left upper submatrix being a $(k+1)\times(k+1)$ matrix (the order of states in $A_{t,0}$ is $(x_1, x_2; 0, 0), (x_1, x_2; 1, 0), \ldots, (x_1, x_2; k-1, 0), (x_1, x_2; 0, 1), \ldots, (x_1, x_2; 0, r-1)$). Moreover, matrices $A^{(1)}_{t,2}, \ldots, A^{(1)}_{t,k-1}$ will be $(k + r) \times (k + r)$ matrices with all their entries 0. Matrix $A^{(1)}_{t,1}$ will have all its entries 0 except for the entry $(k + 1, k + 1)$ which equals p_t. Matrix $A^{(1)}_{t,k}$ will have all its entries 0 except for the entry $(k, k+1)$ which equals p_t. Finally, matrix $A^{(2)}_{t,1}$ will have all its entries 0 except for the entry $(k+r, 1)$ which equals q_t. The appropriate initial probability vector of the Markov chain introduced here is given by $\boldsymbol{\pi}_{0,0} = (1, 0, 0, \ldots, 0)$.

Using Theorem 1 we may readily evaluate the probability mass function of $(X^{(1)}_{n,k}, X^{(2)}_{n,r})$, and exploiting Theorem 2 the associated joint probability generating function will read as follows

$$\varphi_n(z_1, z_2) = \boldsymbol{\pi}_{0,0} \prod_{r=1}^{n} \left(A_{r,0} + z_1 A^{(1)}_{r,1} + z_1^k A^{(1)}_{r,k} + z_2 A^{(2)}_{r,1} \right) \mathbf{1}'.$$

In the case of iid trials with success probabilities p ($p_t = p$, $q_t = q$, for all $t = 1, 2, \ldots$), Theorem 3 yields, after some lengthy but straightforward calculations, that

$$\Phi(z_1, z_2; w) = \sum_{n=0}^{\infty} \varphi_n(z_1, z_2) w^n = \frac{P(z_1, z_2; w)}{Q(z_1, z_2; w)}$$

where

$$
\begin{aligned}
P(z_1,z_2;w) &= 1-(pw)z_1-(pw)^k(1-z_1^k)-(pw)^{k+1}(z_1^k-z_1)\\
&\quad -(qw)^r+(qw)^r(pw)z_1\\
&\quad +(pw)^k(qw)^r(1-z_1^k)+(pw)^{k+1}(qw)^r(z_1^k-z_1)\\
Q(z_1,z_2;w) &= 1-w(1+pz_1)+w^2pz_1-(qw)^rz_2+(qw)^rw(p(1+z_1z_2)+qz_2)\\
&\quad -(qw)^r(pw)wz_1(p+qz_2)+(pw)^k(qw)(1-z_1^k)\\
&\quad +(pw)^{k+1}(qw)(z_1^k-z_1)\\
&\quad -(pw)^k(qw)^r(1-z_2)(1-z_1^k)-(pw)^k(qw)^rw[(qz_2(1-z_1^k)\\
&\quad +p(z_1^k-z_1)(1-z_2)]-(pw)^{k+1}(qw)^{r+1}z_2(z_1^k-z_1).
\end{aligned}
$$

Needless to say $\Phi(z_1,1;w)$ reduces to the double pgf of $X_{n,k}^{(1)}$ (see, Antzoulakos et al., 2003) while $\Phi(1,z_2;w)$ coincides with the double pgf of $X_{n,r}^{(2)}$ (see e.g. Koutras and Alexandrou, 1995).

For the benefit of the practical minded reader we mention that the last expression for $\Phi(z_1,z_2;w)$ may be used to establish a recursive scheme for the joint probability mass function of $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$. More specifically, if we make use of the identity $P(z_1,z_2;w) = (\sum_{n=0}^{\infty}\varphi_n(z_1,z_2)w^n)Q(z_1,z_2;w)$ and pick up the coefficients of w^n $(n=0,1,\ldots)$ in both sides, a set of recurrence relations for $\varphi_n(z_1,z_2)$ will arise. A further manipulation on these, results in a recursive scheme for the joint probability mass function of $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$. This can be used for establishing an effective calculation scheme for the numerical evaluation of the probabilities $P(X_{n,k}^{(1)}=x_1, X_{n,r}^{(2)}=x_2)$. The details are left to the interested reader who may also compare this approach to the alternative calculation option offered by Theorem's 1 outcome. In Fig. 1 the joint probability mass function of $(X_{n,k}^{(1)}, X_{n,r}^{(2)})$ has been pictured for several values of n, k, r and p.

In Table 1 we give values of the correlation coefficient of $X_{n,k}^{(1)}, X_{n,r}^{(2)}$ for several values of n, k, r and p.

The evaluation of the correlation coefficient has been performed using Theorem 4 and results regarding the calculation of second order moments of $X_{n,k}^{(1)}$ and $X_{n,r}^{(2)}$ appeared in Antzoulakos et al. (2003) and Antzoulakos and Chadjicostantinides (2001).

4 Applications

In this section we highlight two potential uses of the distribution described in Sect. 3 in applied research.

The first application comes from the area of reliability theory. A *consecutive-r-out-of-n: F* reliability system consists of n components placed in a line and fails whenever at least r consecutive components fail. In the iid case, the

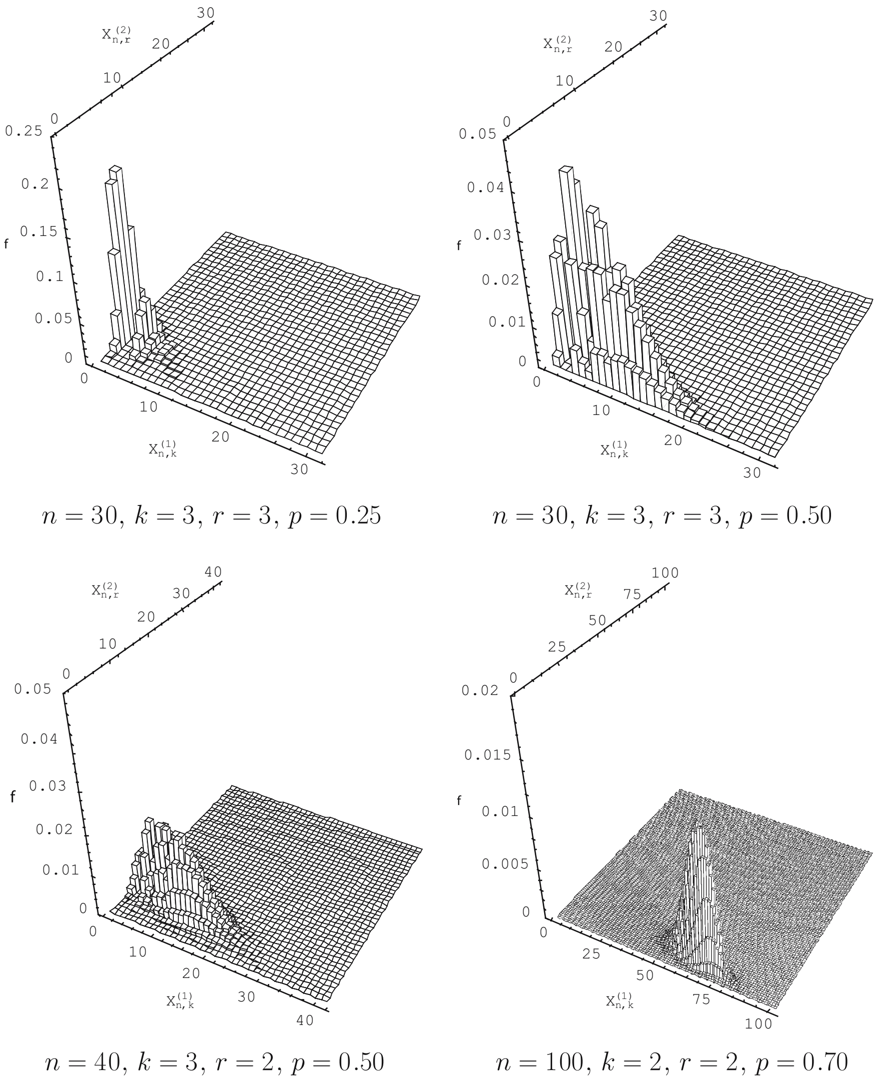

Fig. 1 Distribution of $(X^{(1)}_{n,k}, X^{(2)}_{n,r})$

n components are assumed to work independently of each other and share a common survival probability $p = 1 - q$. The following simple and practical example has been given in the related literature to justify the usefulness of the consecutive-r-out-of-n: F system (see e.g. Chao et al., 1995).

A system of n radar stations is used for transmitting signals from site A to site B. Assume that the n stations are equally spaced between A and B and that each station is able to transmit signals up to a distance of r stations. Apparently, the system becomes non-functional if and only if at least r consecutive radar stations are out of order.

Table 1 Correlation coefficient of $X_{n,k}^{(1)}, X_{n,r}^{(2)}$

n	k	r	p	$\rho(X_{n,k}^{(1)}, X_{n,r}^{(2)})$
30	2	2	0.25	−0.575472
30	2	2	0.50	−0.645813
30	2	3	0.25	−0.486567
30	2	3	0.50	−0.510352
30	2	3	0.75	−0.394063
30	2	3	0.95	−0.119666
30	4	3	0.25	−0.176318
30	4	3	0.50	−0.292866
30	4	3	0.75	−0.260741
30	4	3	0.95	−0.083848
50	4	3	0.50	−0.297753
100	4	3	0.50	−0.305279
100	2	2	0.50	−0.653370
100	2	2	0.95	−0.384958

Let us now assume that, during system's operational periods, the equipment located at big clusters of consecutive working stations is used to produce some additional work. Suppose that this function is carried out only by stations belonging to working clusters of length at least k and that the additional work produced by each cluster is proportional to the cluster length (assume, for example, that each participating radar station produces work of volume 1). Adhering to the notations introduced in the previous section ($Z_i = 1$ if component i is functional and $Z_i = 0$ otherwise), we may easily conclude that the volume of the additional work produced is closely related to the conditional distribution of the random variable $X_{n,k}^{(1)}$ given that $X_{n,r}^{(2)} = 0$. Thus, the probability that the additional work produced is x equals

$$P(X_{n,k}^{(1)} = x \,|\, X_{n,r}^{(2)} = 0) = \frac{P(X_{n,k}^{(1)} = x, X_{n,r}^{(2)} = 0)}{P(X_{n,r}^{(2)} = 0)} = \frac{g_n(x)}{P(X_{n,r}^{(2)} = 0)} \tag{2}$$

while the mean additional work produced is given by

$$E(X_{n,k}^{(1)} \,|\, X_{n,r}^{(2)} = 0) = \frac{1}{P(X_{n,r}^{(2)} = 0)} \sum_{x=0}^{\infty} x g_n(x). \tag{3}$$

The evaluation of the denominator in (2) and (3), which is actually the reliability of a consecutive-r-out-of-n: F system can be easily accomplished by the aid of the recursive scheme (see e.g. Aki and Hirano, 1988)

$$P(X_{n,r}^{(2)} = 0) = P(X_{n-1,r}^{(2)} = 0) - pq^r P(X_{n-r-1,r}^{(2)} = 0), \quad n \geq r+1$$

and the initial conditions

$$P(X_{n,r}^{(2)} = 0) = \begin{cases} 1, & 0 \le n < r \\ 1 - q^r, & n = r. \end{cases}$$

As far as the numerator in the RHS of (2) is concerned, it is not difficult to check that its double generating function is given by

$$\sum_{n=0}^{\infty}\sum_{x=0}^{\infty} g_n(x) z^x w^n = \Phi(z, 0; w) = \frac{P(z, 0; w)}{Q(z, 0; w)}$$

where $\Phi(z, 0; w)$, $P(z, 0; w)$, $Q(z, 0; w)$ are the quantities introduced in the previous section. Restating the last equality in the form

$$P(z, 0; w) = Q(z, 0; w) \sum_{n=0}^{\infty}\sum_{x=0}^{\infty} g_n(x) z^x w^n \tag{4}$$

and taking into account that

$$\begin{aligned} P(z, 0; w) = {} & 1 - (pw)z - (pw)^k(1 - z^k) - (pw)^{k+1}(z^k - z) - (qw)^r \\ & +(qw)^r(pw)z + (pw)^k(qw)^r(1 - z^k) + (pw)^{k+1}(qw)^r(z^k - z) \\ Q(z, 0; w) = {} & 1 - w(1 + pz) + w^2 pz + (qw)^r(pw) \\ & -(qw)^r(pw)^2 z + (pw)^k(qw)(1 - z^k) + (pw)^{k+1}(qw)(z^k - z) \\ & -(pw)^k(qw)^r(1 - z^k) - (pw)^{k+1}(qw)^r(z^k - z) \end{aligned}$$

we may easily derive the following efficient recursive scheme for $g_n(x)$

$$\begin{aligned} g_n(x) = {} & g_{n-1}(x) + p(g_{n-1}(x-1) - g_{n-2}(x-1)) \\ & -pq^r(g_{n-r-1}(x) - pg_{n-r-2}(x-1)) \\ & -p^k q(g_{n-k-1}(x) - g_{n-k-1}(x-k)) \\ & -p^{k+1} q(g_{n-k-2}(x-k) - g_{n-k-2}(x-1)) \\ & +p^k q^r(g_{n-k-r}(x) - g_{n-k-r}(x-k) \\ & +p(g_{n-k-r-1}(x-k) - g_{n-k-r-1}(x-1))), \\ & \qquad n > k + r + 1. \end{aligned}$$

Identity (4) can also be exploited to deduce the initial conditions needed in order to launch the last recursive scheme (the details are left to the reader).

It is worth mentioning that the generating function of the quantity $\mu_n = \sum_{x=0}^{\infty} x g_n(x)$, $n \ge 0$, appearing in formula (3) is given by

$$\sum_{n=0}^{\infty} \mu_n w^n = \left[\frac{\partial \Phi(z, 0; w)}{\partial z}\right]_{z=1} = \frac{(pw)^k(1 - (qw)^r)^2(pw + k(1 - pw))}{(1 - w(1 - p(qw)^r))^2}.$$

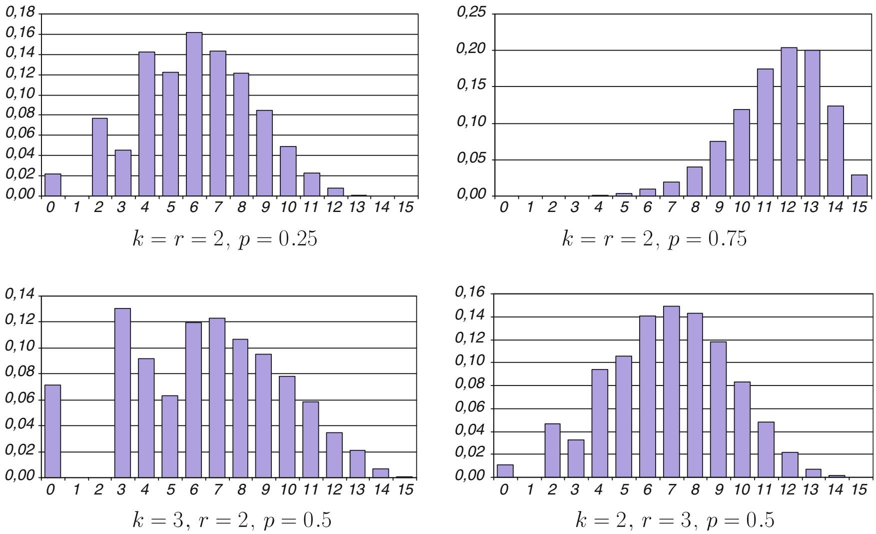

Fig. 2 Conditional distribution of $X_{15,k}^{(1)} \mid X_{15,r}^{(2)} = 0$

This last expression can be exploited to derive the following recurrence relation

$$\mu_n = 2\mu_{n-1} - \mu_{n-2} - 2pq^r(\mu_{n-r-1} - \mu_{n-r-2}) - (pq^r)^2\mu_{n-2r-2}, \quad n \geq k + 2r + 2$$

which offers an efficient scheme for the numerical evaluation of the conditional mean $E(X_{n,k}^{(1)} \mid X_{n,r}^{(2)} = 0)$. In the special case $k = 2$ and $r \geq 4$ the next set of initial conditions should be used before launching the recursive scheme

$$\begin{aligned}
&\mu_0 = \mu_1 = 0, \quad \mu_2 = 2p^2, \quad \mu_3 = p^2(3+q),\\
&\mu_n = 2\mu_{n-1} - \mu_{n-2} - \delta_{r+2}(n)4p^2q^r\\
&\qquad + \delta_{r+3}(n)2p^3q^r - \gamma_{r+4}(n)2pq^r(\mu_{n-r-1} - \mu_{n-r-2})\\
&\qquad + \delta_{2r+2}(n)2(pq^r)^2 - \delta_{2r+3}(n)p^3q^{2r}, \quad 4 \leq n \leq 2r+3
\end{aligned}$$

($\delta_i(j)$ is the Kronecker's delta function, and $\gamma_i(j) = 1$ if $j \geq i$ and $\gamma_i(j) = 0$ if $j < i$).

In Fig. 2 the probability mass function of the conditional distribution of $X_{n,k}^{(1)}$ given that $X_{n,r}^{(2)} = 0$ is given for $n = 15$, and a variety of choices for k, r, and p.

In Table 2 we provide numerical values of $E(X_{n,k}^{(1)} \mid X_{n,r}^{(2)} = 0)$ for several values of n, k, r and p.

Let us next describe an additional example related to educational psychology where the aforementioned probability model can be applied as well. In experimental studies of learning and memory, psychologists usually seek reasonable achievement testing criteria useful for determining the termination of a treatment. One of the older and most familiar criteria in the psychological society is

Table 2 Values of $E(X_{n,k}^{(1)} \mid X_{n,r}^{(2)} = 0)$

n	k	r	$p = 0.25$	$p = 0.50$	$p = 0.75$
15	2	2	6.021598	8.798372	11.539015
15	2	3	3.759799	6.904704	10.635607
15	3	2	3.304513	7.010023	10.312177
15	3	3	1.636391	4.648883	9.253016
15	3	4	1.542217	3.970514	8.969265
15	4	3	0.647951	2.910225	7.756206
15	4	4	0.393088	2.391705	7.464409
30	2	2	12.53640	18.06890	23.40830
30	2	3	7.925930	14.24950	21.59460
30	3	2	7.187790	13.88980	21.32600
30	3	3	3.580250	9.893390	19.18790
30	3	4	2.460000	8.455980	18.59460
30	4	3	1.481480	6.455030	16.52490
30	4	4	0.908257	5.290260	15.90450

Grant's (1946) run criterion which rejects the null hypothesis of no learning if the subject under study provides correct responses in each of a specified number of successive trials. Assume that the probability of correct response in a trial equals q and denote by r the number of consecutive correct responses evidencing the subject's learning achievement. Then the conditional distribution of $X_{n,1}^{(1)}$ given that $X_{n,r}^{(2)} = 0$, provides information on the total number of incorrect responses in n trials given that, up to that trial, the subject has not met Grant's (1946) run criterion requirement. Should we wish to ignore isolated incorrect responses, the mean number of unsuccessful responses in n trials given that the subject has not yet qualified for learning achievement, equals $E(X_{n,2}^{(1)} \mid X_{n,r}^{(2)} = 0)$. Finally, $E(X_{n,k}^{(1)} = x \mid X_{n,r}^{(2)} = 0)$ offers the mean number of unsuccessful responses when blocks of wrong responses of length $1, 2, \ldots, k-1$ are ignored.

References

Agin, M. A., Godbole, A. P. (1992). A new exact runs test for randomness. In C. Page, R. Le Page(Eds.), *Computing science and statistics*, Proceedings of the 22 symposium on the interface (pp. 281–285). Berlin Heidelberg New York: Springer.

Aki, S., Hirano, K. (1988). Some characteristics of the binomial distribution of order k and related distributions. In K. Matusita (Ed.) *Statistical theory and data analysis II,* (pp. 211–222). North-Holland: Elsevier Science.

Antzoulakos, D. L., Chadjicostantinides, S. (2001). Distributions of numbers of success runs of fixed length in Markov dependent trials. *Annals of the Institute of Statistical Mathematics*, *53*, 579–619.

Antzoulakos, D. L., Bersimis, S., Koutras, M. V. (2003). On the distribution of the total number of run lengths. *Annals of the Institute of Statistical Mathematics*, *55*(4), 865–884.

Balakrishnan, N., Koutras, M. V. (2002). *Runs and scans with applications.* New York: Wiley.

Chao, M. T., Fu, J. C., Koutras, M. V. (1995). A survey of the reliability studies of consecutive-k-out-of-n: F systems and its related systems. *IEEE Transactions on Reliability*, *44*, 120–127.

Doi, M., Yamamoto, E. (1998). On the joint distribution of runs in a sequence of multi-state trials. *Statistics and Probability Letters*, *39*, 133–141.

Feller, W. (1968). *An introduction to probability theory and its applications* (Vol.I, 3rd edn) New York: Wiley.

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

Fu, J. C., Koutras, M. V. (1994). Distribution theory of runs: a Markov chain approach. *Journal of the American Statistical Association*, *89*, 1050–1058.

Fu, J. C., Lou, W. Y. W. (2003). *Distribution Theory of Runs and Patterns and Its Application*. Singapore: World Scientific.

Gani, J. (2003). Patterns in sequences of random events. In D.N. Shanbhag, C.R. Rao (Eds.) *Handbook of Statistics*, (Vol. *21*, pp. 227–242).

Glaz, J., Naus, J., Wallenstein, S. (2001). *Scan statistics*. New York: Springer.

Grant, D. (1946). New statistical criteria for learning and problem solution in experiments involving repeated trials. *Psychological Bulletin*, *43*, 272–282.

Guibas, L. J., Odlyzko, A. M. (1981). String overlaps, pattern matching and nontransitive games. *Journal of Combinatorial Theory, Series A*, *30*, 183–208.

Han, Q., Aki, S. (1999). Joint distributions of runs in a sequence of multistate trials. *Annals of the Institute of Statistical Mathematics*, *51*, 419–447.

Koutras, M. V. (1997). Waiting time distributions associated with runs of fixed length in two-state Markov chains. *Annals of the Institute of Statistical Mathematics*, *49*, 123–139.

Koutras, M. V. (2003). Applications of Markov chains to the distribution theory of runs and patterns. In D. N. Shanbhag, C. R. Rao(Eds.), *Handbook of statistics*, (Vol.*21*, pp. 431–472).

Koutras, M. V., Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach. *Annals of the Institute of Statistical Mathematics*, *47*, 743–766.

Leslie, R. T. (1967). Recurrent composite events. *Journal of Applied Probability*, *4*, 34–61.

Lou, W. Y. W. (2003). The exact distribution of the k-tuple statistic for sequence homology. *Statistics and Probability Letters*, *61*, 51–59.

Mood, A. M. (1940). The distribution theory of runs. *Annals of Mathematical Statistics*, *11*, 367–392.

O'Brien, P. C., Dyck, P. J. (1985). A runs test based on run lengths. *Biometrics*, *41*, 237–244.

Rajarshi, M. B.(1974). Success runs in a two-state Markov chain. *Journal of Applied Probability*, *11*, 190–192.