Ann. Inst. Statist. Math.
Vol. 49, No. 1, 123–139 (1997)

WAITING TIME DISTRIBUTIONS ASSOCIATED WITH RUNS OF FIXED LENGTH IN TWO-STATE MARKOV CHAINS*

M. V. KOUTRAS

Department of Mathematics, University of Athens, Panepistemiopolis, 15784, Greece

(Received September 18, 1995; revised February 6, 1996)

Abstract. In the present article a general technique is developed for the evaluation of the exact distribution in a wide class of waiting time problems. As an application the waiting time for the r-th appearance of success runs of specified length in a sequence of outcomes evolving from a first order two-state Markov chain is systematically investigated and asymptotic results are established. Several extensions and generalisations are also discussed.

Key words and phrases: Waiting time distributions, success runs, distributions of order k, negative binomial distributions.

1. Introduction

The origin of problems related to success run waiting times in sequences of Bernoulli trials goes back as far as De Moivre's era. Their widespread applicability in numerous scientific fields (psychology, meteorology, non-parametric statistical inference, quality control etc.) caused a continuous research interest and led to several variations, extensions and generalisations of the original concept and set-up. Two contemporary areas where the application of distribution theory of runs had beneficial influence are reliability theory (cf. Chao *et al.* (1995) for a review) and start-up demonstration tests, Balakrishnan *et al.* (1995).

The classical framework for a run-related problem is the one proposed by Feller (1968). A sequence of independent Bernoulli trials (with two possible outcomes: Success (S) or failure (F)) is generated and the number of *non-overlapping* occurrences of k consecutive successes is counted. A second enumeration scheme is generated by the *overlapping* counting where an uninterrupted sequence of $l \geq k$ successes preceded and followed by a failure accounts for $l - k + 1$ runs. Finally one more enumeration procedure can be initiated by counting a succession of *at least* k S's only once, irrespectively of its actual length. To fix the distinction between the aforementioned enumeration methods we mention by way of example

* This work was supported by the University of Athens Research Foundation under Grant 70/4/2551.

that for $k = 2$, the sequence $FSSSFSSSSSFS$ contains 3, 6, 2 non-overlapping, overlapping and "greater than" runs respectively.

The distributions of success runs of fixed length k have been termed for obvious reasons as *distributions of order k*. For more details we refer to Johnson *et al.* (1992) where a special paragraph is devoted to them.

Although Feller's way of counting simplifies the probabilistic model (which can be easily analysed by renewal theory arguments) it depends on the particular application we are dealing with, which method of enumeration is the most appropriate. So, in reliability theory, the non-overlapping counting is suitable for the definition of a specific structure called m-consecutive-k-out-of-n: F system (see Papastavridis and Koutras (1993) for a review) whereas in molecular biology, DNA sequence matching problems are handled by counting the "coincidence" runs of length at least k, Goldstein (1990).

Another direction for generalising Feller's model is to drop the assumption of independency and consider instead a sequence of trials evolving according to a stationary Markov chain. Run-related problems under Markovian dependence set-ups have been recently studied by Schwager (1983), Aki and Hirano (1993), Hirano and Aki (1993), Mohanty (1994), Uchida and Aki (1995), etc.

In the present paper we conduct a systematic study of the waiting time distribution for the r-th appearance of a success run (non-overlapping, at least, overlapping) of length k in a sequence of Markov dependent trials. After the introduction of the necessary notations and definitions (Section 2), some general results are established in Section 3 which offer universal tools for the study of waiting time distributions. This is accomplished by exploiting the Markov chain imbedding technique introduced recently by Fu and Koutras (1994) and subsequently refined by Koutras and Alexandrou (1995) (see also Fu (1996) where the method is extended for multistate trials and general compound patterns). In Section 4, the exact waiting time distributions for success runs are explored and a representation is established for each of them as sum of independent random variables. Section 5 deals with asymptotic results and finally, Section 6 discusses a number of possible extensions and generalisations of the presented material.

2. Definitions and notations

Let $X_0, X_1, X_2, \ldots$ be a time homogeneous two-state Markov chain with transition probability matrix

$$P = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix}$$

i.e.

$$p_{ij} = P(X_t = j \mid X_{t-1} = i), \qquad t \geq 1, \;\; 0 \leq i, j \leq 1 \tag{2.1}$$

and initial probabilities $p_j = P(X_0 = j)$, $j = 0, 1$. The waiting time for the r-th occurrence of a success run of length k (r, k positive integers) will be denoted by $T_{r,k}^{(a)}$ or simply $T_r^{(a)}$ with the superscript pointing out the enumeration scheme employed; $a = I$ will indicate the non overlapping counting, $a = II$ the "at least" scheme and $a = III$ the overlapping one. For the respective distributions we are going to use the names *type I, II, III Markov Negative Binomial distributions of order k*, in symbols $MNB^{(I)}_{r,k}$, $MNB^{(II)}_{r,k}$, $MNB^{(III)}_{r,k}$.

Let

$$h^{(a)}_r(n) = P(T^{(a)}_r = n), \qquad n \geq 0$$

be the probability distribution function of $T^{(a)}_r$ and

$$H^{(a)}_r(z) = \sum_{n=0}^{\infty} h^{(a)}_r(n) z^n, \qquad r \geq 0$$

$$H^{(a)}(z,w) = \sum_{r=0}^{\infty} H^{(a)}_r(z) w^r$$

its single and double probability generating functions (convention: $H^{(a)}_0(z) = 0$). For the special case $r = 1$ it is clear that $MNB^{(I)}_{1,k} = MNB^{(II)}_{1,k} = MNB^{(III)}_{1,k}$ and the resulting distribution will be referred as *Markov Geometric distribution of order k* (in symbols: MG_k), its probability distribution function and (single) probability generating function being denoted hereafter by $h(n)$, $H(z)$ (instead of $h^{(a)}_1(n)$, $H^{(a)}_1(z)$).

For the benefit of typographic simplicity and when no confusion is likely to arise, formulae and results encompassing all three enumeration schemes will be stated without the use of superscripts.

Evidently, for $p_{i1} = p_1 = p$, $p_{i0} = p_0 = 1 - p$, the model reduces to the iid framework and $MNB^{(I)}_{r,k}$, $MNB^{(III)}_{r,k}$ coincide with the ordinary Negative Binomial distributions of order k (cf. Philippou *et al.* (1983) and Hirano *et al.* (1991)).

3. General results

In a recent paper, 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 imbedding technique. Koutras and Alexandrou (1995) refined this method and expressed the distribution of several run-related statistics in terms of multidimensional binomial type probability vectors; their approach facilitated the establishment of exact formulae for the respective probability generating functions in terms of certain matrices which characterise the enumeration scheme in use. In the sequel, we are going to exploit this approach for constructing general formulae for both the probability distribution function and probability generating function of Markov Negative Binomial distributions of order k.

Before advancing to the main part of this paragraph, we deem necessary to give a brief outline of the aforementioned Markov chain imbedding technique and some of its machinery which is essential in the derivation of our results. For more details we refer to Fu and Koutras (1994) and Koutras and Alexandrou (1995).

A discrete random variable V_n defined on $\{0, 1, \ldots, l_n\}$ (n a non-negative integer) will be called *Markov chain imbeddable Variable of Binomial type* (MVB) if

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

$$\Omega = \bigcup_{v \geq 0} C_v, \qquad C_v = \{c_{v,0}, c_{v,1}, \ldots, c_{v,s-1}\}.$$

2. $P(Y_t \in C_w) \mid Y_{t-1} \in C_v) = 0$ for all $w \neq v, v+1$ and $t \geq 1$.
3. The event $V_n = v$ is equivalent to $Y_n \in C_v$ and therefore the probability distribution function of V_n is given by

$$P(V_n = v) = P(Y_n \in C_v), \qquad v = 0, 1, \ldots, l_n.$$

The distribution of a *MVB* is completely determined by the following three quantities:

- the initial probabilities

$$\pi_v = (P(Y_0 = c_{v,0}), P(Y_0 = c_{v,1}), \ldots, P(Y_0 = c_{v,s-1}))$$

- the *within states* one step transition matrix

$$A_t(v) = (P(Y_t = c_{v,j} \mid Y_{t-1} = c_{v,i}))_{s \times s}$$

- the *between states* one step transition matrix

$$B_t(v) = (P(Y_t = c_{v+1,j} \mid Y_{t-1} = c_{v,i}))_{s \times s}.$$

More specifically, as Koutras and Alexandrou (1995) indicated, if

$$\boldsymbol{f}_t(v) = (P(Y_t = c_{v,0}), P(Y_t = c_{v,1}), \ldots, P(Y_t = c_{v,s-1}))$$

then the next recurrences hold true for all $1 < t \leq n$,

$$\begin{aligned} &\boldsymbol{f}_t(0) = \boldsymbol{f}_{t-1}(0) A_t(0) \\ &\boldsymbol{f}_t(v) = \boldsymbol{f}_{t-1}(v) A_t(v) + \boldsymbol{f}_{t-1}(v-1) B_t(v-1), \qquad 1 \leq v \leq l_n. \end{aligned} \tag{3.1}$$

These relations, in conjunction with the initial conditions $\boldsymbol{f}_0(v) = \pi_v$, $0 \leq v \leq l_n$ offer a very simple computational scheme for the evaluation of the probability distribution function of V_n through the formula

$$P(V_n = v) = \boldsymbol{f}_n(v)\mathbf{1}', \qquad v = 0, 1, \ldots, l_n$$

($\mathbf{1} = (1, 1, \ldots, 1)$ is the row vector of R^s with all its entries being 1).

It is sufficient for our purposes and also of great simplicity (especially for the statement of more compact formulae) to assume that $\pi_v = \mathbf{0}$, $v \geq 1$ and $\pi_0 \mathbf{1}' = 1$; this convention is in fact equivalent to the condition $P(V_0 = 0) = 1$.

Let now T_r, $r \geq 1$ be the waiting time for the r-th occurrence of the event enumerated by V_n, i.e. T_r takes the value n if and only if $V_n = r$ and $V_{n-1} = r - 1$. Therefore the probability distribution function $h_r(n)$ of T_r can be written as

$$\begin{aligned} h_r(n) &= P(Y_n \in C_r, Y_{n-1} \in C_{r-1}) \\ &= \sum_{i=0}^{s-1} P(Y_n \in C_r \mid Y_{n-1} = c_{r-1,i}) P(Y_{n-1} = c_{r-1,i}). \end{aligned}$$

Denoting by $\boldsymbol{e}_i = (0, \dots, 1, \dots, 0)$ the i-th unit vector of R^s and by $\beta_i(n, r)$ the quantity

$$\beta_i(n, r) = \boldsymbol{e}_i B_n(r-1)\mathbf{1}', \qquad 1 \leq i \leq s$$

we may readily verify that $P(Y_{n-1} = c_{r-1,i}) = \boldsymbol{f}_{n-1}(r-1)\boldsymbol{e}'_{i+1}$ and

$$P(Y_n \in C_r \mid Y_{n-1} = c_{r-1,i}) = \sum_{j=1}^{s} \boldsymbol{e}_{i+1} B_n(r-1)\boldsymbol{e}'_j = \beta_{i+1}(n, r).$$

This yields the following interesting expression for the probability distribution function of T_r

$$h_r(n) = \sum_{i=1}^{s} \beta_i(n, r) \boldsymbol{f}_{n-1}(r-1)\boldsymbol{e}'_i. \tag{3.2}$$

It is worth mentioning that the exact distribution of T_r can alternatively be evaluated by making use of Theorem 5.1 in Fu (1996) which in fact covers much broader waiting time problems (multistate trials, general compound patterns); nevertheless, the present approach offers a computationally more efficient scheme and on the other hand facilitates the establishment of general formulae for the respective probability generating functions in the event of homogeneous MVB's (which fortunately is the case for the models we intend to apply our general results to). The next two theorems deal with this situation.

THEOREM 3.1. *If $A_t(v) = A$, $B_t(v) = B$ for all t, $v \geq 0$ then the double generating function of T_r is given by*

$$H(z, w) = wz\boldsymbol{\pi}_0 \sum_{i=1}^{s} \beta_i [I - z(A + wB)]^{-1} \boldsymbol{e}'_i \tag{3.3}$$

where $\beta_i = \boldsymbol{e}_i B \mathbf{1}'$, $1 \leq i \leq s$.

PROOF. A straightforward manipulation over (3.2) reveals that

$$H(z, w) = wz \sum_{i=1}^{s} \beta_i \left(\sum_{n=0}^{\infty} \sum_{r=0}^{l_n} \boldsymbol{f}_n(r) w^r z^n \right) \boldsymbol{e}'_i$$

and the final formula for $H(z,w)$ is readily ascertainable from the identity

$$\sum_{n=0}^{\infty}\sum_{r=0}^{l_n} \boldsymbol{f}_n(r) w^r z^n = \boldsymbol{\pi}_0[I - z(A+wB)]^{-1}$$

which appears in Koutras and Alexandrou (1995).

THEOREM 3.2. *If $A_t(v) = A$, $B_t(v) = B$ for all t, $v \geq 0$ then the probability generating function of T_r can be expressed as*

$$H_r(z) = z^r \boldsymbol{\pi}_0 \sum_{i=1}^{s} \beta_i [(I - zA)^{-1}B]^{r-1}(I - zA)^{-1}\boldsymbol{e}_i', \qquad r \geq 1. \tag{3.4}$$

PROOF. Since

$$I - z(A + wB) = (I - zA)[I - zw(I - zA)^{-1}B]$$

it follows that

$$[I - z(A + wB)]^{-1} = \sum_{j=0}^{\infty}[(I - zA)^{-1}B]^j (I - zA)^{-1}(zw)^j$$

and substituting in (3.3) we can write $H(z,w)$ as

$$H(z,w) = \boldsymbol{\pi}_0 \sum_{i=1}^{s} \beta_i \sum_{j=0}^{\infty}[(I - zA)^{-1}B]^j (I - zA)^{-1}(zw)^{j+1}\boldsymbol{e}_i'$$

or equivalently in the more interesting form

$$H(z,w) = \sum_{r=1}^{\infty} z^r \boldsymbol{\pi}_0 \left(\sum_{i=1}^{s} \beta_i [(I - zA)^{-1}B]^{r-1}(I - zA)^{-1}\boldsymbol{e}_i' \right) w^r$$

which manifestly yields the desired result.

Applying Theorem 3.2 for the special case $r = 1$ we get the following expression for the probability generating function $H_1(z) \equiv H(z)$ of the first occurrence time T_1

$$H(z) = z\boldsymbol{\pi}_0 \sum_{i=1}^{s} \beta_i (I - zA)^{-1}\boldsymbol{e}_i'. \tag{3.5}$$

4. The exact distribution of run waiting times

In this section we are going to conduct a detailed study of the waiting time distribution for the r-th occurrence of a success run of length k in a sequence of Markov dependent trials $X_1, X_2, \ldots$ defined by (2.1). Each one of the three enumeration schemes (non-overlapping, at least, overlapping) is treated separately and besides the probability distribution function of the corresponding variable a representation as a sum of independent variables is established as well.

a. *Non-overlapping success runs*

Following Koutras and Alexandrou (1995), we define $C_v = \{(v,i) : 0 \le i \le k-1\}$ for all $v = 0, 1, \ldots, [n/k]$ ($l_n = [n/k]$, $s = k$) and introduce a proper Markov chain $\{Y_t, t \ge 0\}$ as follows: $Y_t = (v,i)$ if and only if in the sequence of outcomes leading to the t-th trial (say $SFSSF\ldots F\overbrace{SS\ldots S}^{m}$), there exist v non-overlapping success runs and m trailing successes with $m = i (\mathrm{mod}\, k)$. Clearly, the resulting Markov chain is homogeneous, $\boldsymbol{\pi}_0 = (p_0, p_1, 0, \ldots, 0) = p_0 \mathbf{e}_1 + p_1 \mathbf{e}_2$ and the non vanishing entries of matrices $A_t(v) = A = (a_{ij})_{s\times s}$, $B_t(v) = B = (b_{ij})_{s\times s}$ are $a_{11} = p_{00}$, $a_{12} = p_{01}$, $a_{k1} = p_{10}$,

$$a_{ij} = \begin{cases} p_{10} & \text{if} \quad j = 1 \\ p_{11} & \text{if} \quad j = i+1 \end{cases} \qquad 2 \le i \le k-1 \tag{4.1}$$

and $b_{k1} = p_{11}$. Thus

$$\beta_i = \mathbf{e}_i B \mathbf{1}' = \begin{cases} 0 & \text{if} \quad 1 \le i \le k-1 \\ p_{11} & \text{if} \quad i = k \end{cases}$$

and the evaluation of the double generating function of $T_r^{(I)}$ may be easily performed through (3.3); clearly the only quantities we need to this end are the determinant and the $(k,1)$, $(k,2)$ minors of the matrix

$$I - z(A + wB) = \begin{bmatrix} 1 - p_{00}z & -p_{01}z & 0 & \cdot & 0 & 0 \\ -p_{10}z & 1 & -p_{11}z & \cdot & 0 & 0 \\ -p_{10}z & 0 & 1 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -p_{10}z & 0 & 0 & \cdot & 1 & -p_{11}z \\ -(p_{10} + wp_{11})z & 0 & 0 & \cdot & 0 & 1 \end{bmatrix}_{k\times k}.$$

After some routine calculations we get

$$H^{(I)}(z,w) = \frac{w[p_1 + (p_0 p_{01} - p_1 p_{00})z](p_{11}z)^{k-1}}{1 - p_{00}z - p_{01}p_{10}z^2 \sum_{i=2}^{k} (p_{11}z)^{i-2} - wp_{01}p_{11}^{k-1}z^k}$$

which, on introducing the notation

$$\begin{aligned} P(z) &= p_1 + (p_0 p_{01} - p_1 p_{00})z \\ Q(z) &= 1 - p_{00}z - p_{01}p_{10}z^2 \sum_{i=2}^{k} (p_{11}z)^{i-2} \end{aligned} \tag{4.2}$$

takes on the more appealing form

$$H^{(I)}(z,w) = \frac{wP(z)(p_{11}z)^{k-1}}{Q(z) - wp_{01}p_{11}^{k-1}z^k}. \tag{4.3}$$

The generating function of the random variable $T = T_1$ for the first occurrence of a success run of length k will be given by

$$H(z) = \frac{P(z)(p_{11}z)^{k-1}}{Q(z)}, \tag{4.4}$$

as may be easily checked either by (3.5) or more quickly by substituting (4.3) in the obvious formula

$$H(z) = \left[\frac{1}{w}H^{(I)}(z,w)\right]_{w=0}.$$

Another random variable which will be proved very useful in expressing the Markov Negative Binomial distributions as sums of independent random variables is the waiting time T^* for the first occurrence of a success run in a sequence of Markov dependent trials with transition probabilities (2.1) and initial conditions $P(X_0 = 0) = 1$, $P(X_0 = 1) = 0$. Its probability generating function is immediately derived from (4.4) (by setting $p_0 = 1$, $p_1 = 0$) as

$$H^*(z) = \frac{(p_{01}z)(p_{11}z)^{k-1}}{Q(z)}. \tag{4.5}$$

It is noteworthy that the numerical evaluation of both T and T^*'s probability distribution function can be achieved fairly easy through the recurrences implied by (4.4) and (4.5). More specifically, multiplying both sides of (4.4) by $Q(z)$ and equating the coefficients of z^n we may deduce the following recursive scheme

$$h(n) = p_{00}h(n-1) + \sum_{i=2}^{k} p_{01}p_{10}p_{11}^{i-2}h(n-i), \qquad n > k \tag{4.6}$$

with initial conditions

$$h(n) = \begin{cases} 0 & \text{if} \quad 0 \le n < k-1 \\ p_1 p_{11}^{k-1} & \text{if} \quad n = k-1 \\ p_0 p_{01} p_{11}^{k} & \text{if} \quad n = k. \end{cases}$$

Equally, $h^*(n) = P(T^* = n)$ obeys exactly the same recurrence for $n > k$ while the initial conditions reduce to

$$h(n) = \begin{cases} 0 & \text{if} \quad 0 \le n < k \\ p_{01} p_{11}^{k-1} & \text{if} \quad n = k. \end{cases}$$

The means of T and T^* are

$$\begin{aligned} \mu = E(T) &= \frac{(p_{01}+p_{10}) - (p_{01}+p_1 p_{10})p_{11}^{k-1}}{p_{01}p_{10}p_{11}^{k-1}} \\ \mu^* = E(T^*) &= \frac{(p_{01}+p_{10}) - p_{01}p_{11}^{k-1}}{p_{01}p_{10}p_{11}^{k-1}} \end{aligned} \tag{4.7}$$

as one can readily verify by evaluating the first derivatives of $H(z)$, $H^*(z)$ at $z = 1$. The higher order moments can also be deduced from $H(z)$, but their expressions are not very attractive. Notice that, direct manipulations on (4.6) yield some appealing recurrence relations for the raw (about zero) moments of T and T^*; the details are left to the reader.

We are now ready to show that T_r can be decomposed as a sum of r independent waiting time random variables.

THEOREM 4.1. *If T_j^*, $1 \leq j \leq r-1$ are independent duplicates of T^* (with probability generating function* (4.5)) *which are also independent of T (with probability generating function* (4.4)) *then*

$$T_r^{(I)} \stackrel{\mathrm{d}}{=} T + \sum_{j=1}^{r-1} T_j^*. \tag{4.8}$$

PROOF. Expanding (4.3) in a Taylor series around $w = 0$ and considering the coefficient of w^r we obtain the following expression for the probability generating function of $T_r^{(I)}$ (an alternative derivation by Theorem 3.2 is also feasible but since we already have the double generating function (4.3) the power series expansion method is considerably simpler)

$$H_r^{(I)}(z) = \frac{P(z)(p_{11}z)^{k-1}}{Q(z)} \left[\frac{(p_{01}z)(p_{11}z)^{k-1}}{Q(z)}\right]^{r-1}.$$

Now, in view of (4.4) and (4.5) we may write

$$H_r^{(I)}(z) = H(z)[H^*(z)]^{r-1} \tag{4.9}$$

which manifestly implies the representation (4.8).

The generating function formulae given above are of course consistent with analogous results published by Aki and Hirano (1993), Mohanty (1994) and Uchida and Aki (1995). Some slight discordances in the final expressions reflect the different set-up used by them for the evolution of the Markov dependent sequence $X_0, X_1, X_2, \ldots$.

Representation (4.8) does merit a special discussion. It is well known that the usual Negative Binomial distribution is the r-th convolution of the ordinary geometric distribution by itself. On the other hand, the convolution of r iid Geometric distributions of order k gives genesis to the Negative Binomial distribution of order k, Philippou *et al.* (1983), Philippou (1984). Formula (4.8) indicates that the relation between $MNB_{r,k}^{(I)}$ and MG_k is almost so but not quite. In this case, instead of r uniform renewals we have at first a slight irregularity (actually generated by the initial trial X_0) until the first success run of length k is formed; from this point on a completely regenerative procedure commences which becomes anew every time we come up with an additional success run.

b. *Success runs of length at least k*

This case was tackled by Fu and Koutras (1994) by incorporating in Ω an additional "waiting state" which is entered upon the completion of a success run of length k and is not abandoned until a failure shows up (which signals the termination of an already counted run). As Koutras and Alexandrou (1995) indicated, we now have $l_n = [(n+1)/(k+1)]$, $s = k+1$ and the only non zero elements of the transition matrices are given by (4.1) and

$$a_{11} = p_{00}, \qquad a_{12} = p_{01}, \qquad a_{k1} = a_{k+1,1} = p_{10}, \qquad a_{k+1,k+1} = p_{11},$$
$$b_{k,k+1} = p_{11}.$$

Accordingly

$$\beta_i = \boldsymbol{e}_i B \boldsymbol{1}' = \begin{cases} 0 & \text{if } \; 1 \le i \le k-1 \quad \text{or} \quad i = k+1 \\ p_{11} & \text{if } \; i = k \end{cases}$$

and evaluating the determinant and the $(k,1)$, $(k,2)$ minors of the matrix

$$I - z(A + wB) = \begin{bmatrix} 1 - p_{00}z & -p_{01}z & 0 & \cdot & 0 & 0 \\ -p_{10}z & 1 & -p_{11}z & \cdot & 0 & 0 \\ -p_{10}z & 0 & 1 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -p_{10}z & 0 & 0 & \cdot & 1 & -wp_{11}z \\ -p_{10}z & 0 & 0 & \cdot & 0 & 1 - p_{11}z \end{bmatrix}_{(k+1)\times(k+1)},$$

we deduce, in lieu of (3.3)

$$H^{(II)}(z,w) = \frac{w(1 - p_{11}z)P(z)(p_{11}z)^{k-1}}{(1 - p_{11}z)Q(z) - wp_{01}p_{10}p_{11}^{k-1}z^{k+1}} \tag{4.10}$$

($P(z)$, $Q(z)$ are as in (4.2)). We are now ready to establish a representation of $T_r^{(II)}$ as a sum of independent variables, i.e. Theorem's 4.1 analogue for success runs of length at least k.

THEOREM 4.2. *Let T, T_j^*, $1 \le j \le r-1$ be as in Theorem 4.1 and T_j^{**}, $1 \le j \le r-1$ be ordinary geometric random variables (defined in a sequence of iid Bernoulli trials) with success probabilities p_{11}, i.e.*

$$P(T_j^{**} = n) = p_{10}p_{11}^{n-1}, \qquad n \ge 1.$$

If all these variables are mutually independent then

$$T_r^{(II)} \stackrel{\mathrm{d}}{=} T + \sum_{j=1}^{r-1}(T_j^* + T_j^{**}). \tag{4.11}$$

PROOF. Writing (4.10) as

$$H^{(II)}(z,w) = w\frac{P(z)(p_{11}z)^{k-1}}{Q(z)}\left[1 - w\frac{p_{10}z}{1-p_{11}z}\cdot\frac{(p_{01}z)(p_{11}z)^{k-1}}{Q(z)}\right]^{-1}$$

and expanding in a Taylor series around $w = 0$, we get the following expression for the probability generating function of $T_r^{(II)}$

$$H_r^{(II)}(z) = \frac{P(z)(p_{11}z)^{k-1}}{Q(z)}\left[\frac{p_{10}z}{1-p_{11}z}\cdot\frac{(p_{01}z)(p_{11}z)^{k-1}}{Q(z)}\right]^{r-1}.$$

Taking into account (4.4), (4.5) and the fact that the probability generating function of the T_j^{**}'s is $G(z) = p_{10}z/(1-p_{11}z)$ we may write

$$H_r^{(II)}(z) = H(z)[G(z)H^*(z)]^{r-1}. \tag{4.12}$$

Representation (4.11) results immediately from the established equality between the generating functions of $T_r^{(II)}$ and $T + \sum_{j=1}^{r-1}(T_j^* + T_j^{**})$.

The rationale of (4.11) is easily elucidated by rewriting it as

$$T_r^{(II)} \stackrel{\mathrm{d}}{=} (T + T_1^*) + \sum_{j=1}^{r-2}(T_j^{**} + T_{j+1}^*) + T_{r-1}^{**}.$$

The random variable $T + T_1^*$ is the waiting time for the first failure *after* the completion of the first success run of length k. From this point, a regenerative phenomenon is activated which is renewed upon the appearance of the first failure *after* a success run of length k; the associated variables $T_j^{**} + T_{j+1}^*$, $1 \le j \le r-2$, are of the same type as $T + T_1^*$ apart from the irregularity of the latter caused by the presence of X_0. The last random variable T_{r-1}^{**} accounts for the formulation of the r-th success run of length k.

c. *Overlapping success runs*

With the imbedding technique introduced by Fu and Koutras (1994) in mind (see also Koutras and Alexandrou (1995)) we now have $l_n = n - k + 1$, $s = k + 1$ whereas the only non zero entries of the matrices A, B are given by (4.1) and

$$a_{11} = p_{00}, \qquad a_{12} = p_{01}, \qquad a_{k1} = a_{k+1,1} = p_{10}, \qquad b_{k,k+1} = b_{k+1,k+1} = p_{11}.$$

Hence

$$\beta_i = \boldsymbol{e}_i B\mathbf{1}' = \begin{cases} 0 & \text{if} \quad 1 \le i \le k-1 \\ p_{11} & \text{if} \quad i = k \quad \text{or} \quad i = k+1, \end{cases}$$

$$I - z(A + wB) = \begin{bmatrix} 1 - p_{00}z & -p_{01}z & 0 & \cdot & 0 & 0 \\ -p_{10}z & 1 & -p_{11}z & \cdot & 0 & 0 \\ p_{10}z & 0 & 1 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -p_{10}z & 0 & 0 & \cdot & 1 & -wp_{11}z \\ -p_{10}z & 0 & 0 & \cdot & 0 & 1 - wp_{11}z \end{bmatrix}_{(k+1)\times(k+1)},$$

and working in a similar fashion as before we obtain the next form for the double generating function of $T_r^{(III)}$

$$H^{(III)}(z, w) = \frac{wP(z)(p_{11}z)^{k-1}}{(1 - wp_{11}z)Q(z) - wp_{01}p_{10}p_{11}^{k-1}z^{k+1}} \tag{4.13}$$

($P(z)$, $Q(z)$ are as in (4.2)). Regarding the representation of $T_r^{(III)}$ as a sum of independent variables we have the following.

THEOREM 4.3. *Let T, T_j^*, $1 \le j \le r-1$ be as in Theorem 4.1 and W_j, $1 \le j \le r-1$ be independent Bernoulli variables with success probabilities p_{11}. If T_j^*, W_j are mutually independent and*

$$W_j^* = \begin{cases} 1 & \text{if} \quad W_j = 1 \\ 1 + T_j^* & \text{if} \quad W_j = 0 \end{cases}$$

then T, $W_1^*, \ldots, W_{r-1}^*$ are independent and

$$T_r^{(III)} \overset{\mathrm{d}}{=} T + \sum_{j=1}^{r-1} W_j^*. \tag{4.14}$$

PROOF. It is not difficult to verify (either through Theorem 3.2 or better by writing (4.13) as

$$H^{(III)}(z, w) = w\frac{P(z)(p_{11}z)^{k-1}}{Q(z)} \cdot \left\{1 - w\left[(p_{11}z) + (p_{10}z)\frac{(p_{01}z)(p_{11}z)^{k-1}}{Q(z)}\right]\right\}^{-1}$$

and expanding in power series around $w = 0$) that the probability generating function of $T_r^{(III)}$ takes the form

$$H_r^{(III)}(z) = H(z)\left[(p_{11}z) + (p_{10}z)H^*(z)\right]^{r-1}. \tag{4.15}$$

Due to the definitions, the probability generating function of W_j^* is

$$G^*(z) = (p_{11}z) + (p_{10}z)H^*(z)$$

whereas the independency of W_j, T_j^*, $1 \leq j \leq r-1$ and T guarranties the independency of W_j^*, $1 \leq j \leq r-1$ and T. Accordingly, the probability generating function of the sum $T + \sum_{j=1}^{r-1} W_j^*$ coincides with $H_r^{(III)}(z)$; hence the result.

In the special case of iid Bernoulli trials, the results of the present paragraph for overlapping success runs reduce to the ones derived by Hirano *et al.* (1991).

Closing, we mention that Uchida and Aki (1995) have also given formulae for the probability generating functions (4.12), (4.15) with a slight discrepancy in the final expression, due to the different setup used there. Their derivation was based on a completely different technique: the method of generalised (conditional) probability generating functions.

5. Asymptotic behaviour

The limiting behaviour of Markov Negative Binomial distributions as $r \to \infty$ is closely related to the class of Poisson-stopped sum distributions, that is to say distributions of the sum of Poisson number of iid random variables. The definite reference on Poisson-stopped sums is Johnson *et al.* (1992). For the needs of our presentation we can stick to the special case of variables with finite support, thereof obtaining a family of distributions with probability generating function

$$\psi(z; \lambda_1, \lambda_2, \ldots, \lambda_m) = \exp(-\sum_{i=1}^{m} \lambda_i + \sum_{i=1}^{m} \lambda_i z^i)$$

and distribution function of the form

$$f(n; \lambda_1, \lambda_2, \ldots, \lambda_m) = \sum \exp(\sum_{i=1}^{m} \lambda_i) \frac{\prod_{j=1}^{m} \lambda_j^{y_j}}{\prod_{j=1}^{m} y_j!}$$

where the first summation is performed over all non-negative integers $y_1, y_2, \ldots, y_m$ such that $\sum_{j=1}^{m} j y_j = n$. Following Aki (1985) (see also Aki *et al.* (1984)) we shall refer to these distributions as *extended Poisson distributions* of order m with parameters $\lambda_1, \lambda_2, \ldots, \lambda_m$.

THEOREM 5.1. *If* $\lim_{r\to\infty} r p_{00} = \lambda > 0$ *and* $\lim_{r\to\infty} r p_{10} = \mu > 0$ *then the asymptotic distribution of* $T_r^{(I)} - rk + 1$ *is a mixture of an extended Poisson distribution of order* k *with parameters*

$$\lambda_i = \begin{cases} \lambda & \text{if} \quad i = 1 \\ \mu & \text{if} \quad 2 \leq i \leq k \end{cases}$$

and a shifted duplicate of it; more precisely

$$\lim_{r\to\infty} P(T_r^{(I)} - rk + 1 = n) = p_1 f(n; \lambda, \mu, \ldots, \mu) + p_0 f(n-1; \lambda, \mu, \ldots, \mu).$$

PROOF. Evidently

$$\lim_{r\to\infty} P(z) = p_1 + p_0 z, \qquad \lim_{r\to\infty} r(1 - Q(z)) = \lambda z + \mu \sum_{i=2}^{k} z^i$$

and therefore

$$\lim_{r\to\infty} \frac{zH(z)}{H^*(z)} = \lim_{r\to\infty} \frac{P(z)}{p_{01}} = p_1 + p_0 z,$$

$$\lim_{r\to\infty} (z^{-k} H^*(z))^r = \lim_{r\to\infty} \frac{(1-p_{00})^r[(1-p_{10})^r]^{k-1}}{Q^r(z)}$$

$$= \exp\left\{-\lambda(1-z) - \mu \sum_{i=2}^{k}(1-z^i)\right\}.$$

The probability generating function of the shifted random variable $T_r^{(I)} - rk + 1$, by virtue of (4.9) can be expressed as

$$z^{-rk+1} H_r^{(I)}(z) = \frac{zH(z)}{H^*(z)}[z^{-k}H^*(z)]^r$$

and taking the limit as $r \to \infty$ we get

$$\lim_{r\to\infty} z^{-rk+1} H_r^{(I)}(z) = (p_1 + p_0 z) \exp\left\{-\lambda(1-z) - \mu \sum_{i=2}^{k}(1-z^i)\right\}$$

$$= (p_1 + p_0 z)\psi(z; \lambda, \mu, \ldots, \mu).$$

This completes the proof.

THEOREM 5.2. *If* $\lim_{r\to\infty} rp_{00} = \lambda > 0$ *and* $\lim_{r\to\infty} rp_{10} = \mu > 0$ *then the asymptotic distribution of* $T_r^{(III)} - r - k + 2$ *is a mixture of an extended Poisson distribution of order* k *with parameters* $\lambda_k = \mu$, $\lambda_i = 0$, $i \neq k$ *and a shifted duplicate of it, the mixing parameters being* p_1 *and* p_0.

PROOF. Observe first that, on using (4.15), the probability distribution function of $T_r^{(III)} - r - k + 2$ takes the form

$$z^{-r-k+2} H_r^{(III)}(z) = \frac{H(z)}{z^{k-1}} \cdot [1 - p_{10}(1 - H^*(z))]^{r-1}.$$

But

$$\lim_{r\to\infty} \frac{H(z)}{z^{k-1}} = \lim_{r\to\infty} \frac{p_{11}^{k-1} P(z)}{Q(z)} = p_1 + p_0 z,$$

and since

$$\lim_{r\to\infty} H^*(z) = z^k$$

we also have

$$\lim_{r\to\infty} rp_{10}(1-H^*(z)) = \mu(1-z^k).$$

Combining all the aforementioned relations we deduce the limiting expression

$$\lim_{r\to\infty} z^{-r-k+2}H_r^{(III)}(z) = (p_1+p_0z)\exp\left\{-\mu(1-z^k)\right\}$$

and the proof is completed.

Theorems 5.1 and 5.2 generalise the results given by Philippou *et al.* (1983) and Hirano *et al.* (1991) for the case of iid Bernoulli trials.

It is noteworthy that, as one may easily verify through the probability generating function (4.12), the asymptotic distribution of $T_r^{(II)}$ as $r\to\infty$ (after being shifted to the support $\{0,1,\ldots\}$) is degenerate with its mass placed to infinity.

A further point of interest is that the representations of Theorems 4.1 and 4.3 can be used for obtaining some simple approximations to $NB_{r,k}^{(I)}$, $NB_{r,k}^{(III)}$ as $r\to\infty$. This is accomplished by employing the central limit theorem on the differences $T_r^{(I)}-T$, $T_r^{(III)}-T$ which, as sums of $r-1$ iid variables can be approximated satisfactorily by a proper Normal distribution (we recall that the numerical evaluation of the distribution of T is easily acquired by (4.6)). Note also that, by virtue of Theorems 4.1, 4.3 we have

$$E(T_r^{(I)}-T) = (r-1)E(T_j^*) = (r-1)\mu^*$$
$$E(T_r^{(III)}-T) = (r-1)E(W_j^*) = (r-1)[p_{11}+p_{10}(1+\mu^*)].$$

Analogous formulae hold true for the variances as well, but the final expressions are rather cumbersome.

6. Conclusion-generalisations

One of the main advantages of the Markov chain approach established in Section 3 is that it survives in much more broader framework than the one used so far. For example, should the observed binary sequence involve non identical trials, the exact distribution of $T_r^{(a)}$'s could be easily captured through (3.1), (3.2) after some trivial modifications in the elements of matrices $A_t(v)$, $B_t(v)$; as a matter of fact, their basic form stays unaltered, but their non-zero entries now depend on the index v.

Recently, Aki *et al.* (1996) considered a class of waiting time problems (for first run occurrences) in sequences of second order Markov dependent trials with

$$p_{xy} = P(X_i=1 \mid X_{i-1}=y, X_{i-2}=x)$$
$$q_{xy} = P(X_i=0 \mid X_{i-1}=y, X_{i-2}=x)$$

$(x,y\in\{0,1\})$. It is clear that the general machinery of Section 3 is applicable under these assumptions as well. More precisely, in order to study the distribution of $T_r^{(\alpha)}$, it suffices to consider exactly the same state and space definitions as the ones used in Section 4, and replace the entries of matrix A by

$$a_{11} = q_{00}, \quad a_{12} = p_{00}, \quad a_{21} = q_{01}, \quad a_{23} = p_{01},$$
$$a_{ij} = \begin{cases} q_{11} & \text{if} \quad j = 1 \\ p_{11} & \text{if} \quad j = i + 1 \end{cases} \quad 3 \le i \le k - 1.$$

(The rest non zero elements of A and B depend on the enumeration scheme employed each time; their specification is left to the reader.) The derivation of double and single probability distribution functions is, in view of Theorems 3.1, 3.2, a matter of may be lengthy but straightforward algebraic calculations on the resulting matrix $I - z(A + wB)$. It is worth noticing that with a further modification of A and B's non zero entries, we could effortless accommodate a higher order Markov model.

Another possible variation of the basic set-up is created by placing the outcomes of the trials in a circular (instead of linear) arrangement. The analysis of such models may be performed then by combining the ideas of Koutras *et al.* (1994, 1995) with the vector oriented approach of Section 3.

Finally we mention that our techniques could be routinely extended for the study of waiting time variables arising from a sequence of trials with more than two outcomes and more general (composite) patterns (cf. Schwager (1983) and Fu (1996)).

Acknowledgements

This work was done while the author was visiting Queen Mary and Westfield College (University of London). He would like to express his gratitude to the host institution.

References

Aki, S. (1985). Discrete distributions of order k on a binary sequence, *Ann. Inst. Statist. Math.*, **37**, 205–224.

Aki, S. and Hirano, K. (1993). Discrete distributions related to succession events in a two-state Markov chain, *Statistical Science and Data Analysis* (eds. K. Matusita, M. L. Puri and T. Hayakawa), 467–474, VSP International Science Publishers, Zeist.

Aki, S., Kuboki H. and Hirano, K. (1984). On discrete distributions of order k, *Ann. Inst. Statist. Math.*, **36**, 431–440.

Aki, S., Balakrishnan, N. and Mohanty, S. G. (1996). Sooner and later waiting time problems for success and failure runs in higher order Markov dependent trials, *Ann. Inst. Statist. Math.* (to appear).

Balakrishnan, N., Viveros, R. and Balasubramanian, K. (1995). Start-up demonstration tests under correlation and corrective action, *Naval Res. Logist.*, **42**, 1271–1276.

Chao, M. T., Fu, J. C. and 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.

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* (to appear).

Fu, J. C. and Koutras, M. V. (1994). Distribution theory of runs: a Markov chain approach, *J. Amer. Statist. Assoc.*, **89**, 1050–1058.

Goldstein, L. (1990). Poisson approximation in DNA sequence matching, *Comm. Statist. Theory Methods*, **19**, 4167–4179.

Hirano, K. and Aki, S. (1993). On the number of occurrences of success runs of specified length in a two-state Markov chain, *Statist. Sinica*, **3**, 313–320.

Hirano, K., Aki, S., Kashiwagi, N. and Kuboki, H. (1991). On Ling's binomial and negative binomial distributions of order k, *Statist. Probab. Lett.*, **11**, 503–509.

Johnson, N. L., Kotz, S. and Kemp, A. W. (1992). *Univariate Discrete Distributions*, Wiley, New York.

Koutras, M. V. and Alexandrou, V. A. (1995). Runs, scans and urn model distributions: A unified Markov chain approach, *Ann. Inst. Statist. Math.*, **47**, 743–766.

Koutras, M. V., Papadopoulos, G. K. and Papastavridis, S. G. (1994). Circular overlapping success runs, *Runs and Patterns in Probability* (eds. A. P. Godbole and S. G. Papastavridis), 287–305, Kluwer, Dordrecht.

Koutras, M. V., Papadopoulos, G. K. and Papastavridis, S. G. (1995). Runs on a circle, *J. Appl. Probab.*, **32**, 396–404.

Mohanty, S. G. (1994). Success runs of length-k in Markov dependent trials, *Ann. Inst. Statist. Math.*, **46**, 777–796.

Papastavridis, S. G. and Koutras, M. V. (1993). Consecutive-k-out-of-n systems, *New Trends in System Reliability Evaluation* (ed. K. B. Misra), 228–248, Elsevier.

Philippou, A. N. (1984). The negative binomial distribution of order k and some of its properties, *Biometrical J.*, **26**, 789–794.

Philippou, A. N., Georghiou, C. and Philippou, G. N. (1983). A generalized geometric distribution and some of its properties, *Statist. Probab. Lett.*, **1**, 171–175.

Schwager, S. (1983). Run probabilities in sequences of Markov dependent trials, *J. Amer. Statist. Assoc.*, **78**, 168–175.

Uchida, M. and Aki, S. (1995). Sooner and later waiting time problems in a two-state Markov chain, *Ann. Inst. Statist. Math.*, **47**, 415–433.