Ann. Inst. Statist. Math.
Vol. 55, No. 4, 865–884 (2003)

ON THE DISTRIBUTION OF THE TOTAL NUMBER OF RUN LENGTHS

D. L. ANTZOULAKOS, S. BERSIMIS* AND M. V. KOUTRAS*

Department of Statistics and Insurance Science, University of Piraeus, Karaoli & Dimitriou 80, Piraeus 18534, Greece

(Received September 9, 2002; revised December 16, 2002)

Abstract. In the present paper, we study the distribution of a statistic utilizing the runs length of "reasonably long" series of alike elements (success runs) in a sequence of binary trials. More specifically, we are looking at the sum of exact lengths of subsequences (strings) consisting of k or more consecutive successes (k is a given positive integer). The investigation of the statistic of interest is accomplished by exploiting an appropriate generalization of the Markov chain embedding technique introduced by Fu and Koutras (1994, *J. Amer. Statist. Assoc.*, **89**, 1050–1058) and Koutras and Alexandrou (1995, *Ann. Inst. Statist. Math.*, **47**, 743–766). In addition, we explore the conditional distribution of the same statistic, given the number of successes and establish statistical tests for the detection of the null hypothesis of randomness versus the alternative hypothesis of systematic clustering of successes in a sequence of binary outcomes.

Key words and phrases: Success runs, run lengths, Markov chains, Markov chain embeddable variable of polynomial type, randomness tests.

1. Introduction

In the analysis of experimental trials whose outcomes can be classified into two exclusive categories, a question that comes in naturally is whether reasonable criteria providing evidence of clustering of any of the two categories could be established. These criteria could then be used to detect changes in the underlying process which generates the series of outcomes. Many commonly used criteria for the statistical analysis of such phenomena involve the concept of *runs* i.e. uninterrupted sequences of alike elements bordered at each end by other types of elements or by the beginning or by the end of the complete sequence. For example, many quality control plans base the acceptance/rejection of the sample lot on the occurrence of prolonged sequences of successive working/failed components, Wolfowitz (1943), Balakrishnan *et al.* (1993). For a mechanical engineer performing a start-up test for a new machine, it is reasonable to couch his decision (accepting the machine or rejecting it) on the number of consecutive successful or unsuccessful attempted start-ups, Hahn and Gage (1983), Viveros and Balakrishnan (1993), Balakrishnan *et al.* (1995, 1997). The same model, in the context of reliability, leads to the well known consecutive-k-out-of-n: F system and its variations (for a review refer to Chao *et al.* (1995)). Finally, an additional interesting application of the concept

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

of runs comes from the area of non-parametric runs tests, Gibbons and Chakraborti (1992), Agin and Godbole (1992). In this case, the interest is focused on the conditional distribution of runs or equivalently, on runs defined in a sequence of outcomes of prespecified composition.

In the traditional runs/patterns literature, the criteria used take into account the *number* of runs/patterns observed in the experimental sequence or the *number* of runs of specified length or the waiting time for the occurrence of a prespecified *number* of runs. The distributions of the number of fixed size success runs and the associated waiting time distributions have been termed in the statistical bibliography as *distributions of order k* and have been extensively studied by Philippou and Makri (1986), Aki and Hirano (1988), Godbole (1990, 1992), Hirano *et al.* (1991), Hirano and Aki (1993) etc. For a detailed and systematic exposition of the distribution theory of runs the interested reader may wish to consult the recent monograph by Balakrishnan and Koutras (2002).

None of the aforementioned criteria makes use of the exact *length of runs* appearing in the outcome sequence. Agin and Godbole (1992), aiming at the development of non-parametric randomness tests, suggested using run lengths of variable sizes (see also Koutras and Alexandrou (1997)). Of a similar flavour can be considered the approach set in O'Brien and Dyck (1985), where a test based on longest runs is investigated. Motivated by those works, we proceed here to the investigation of a statistic utilizing the runs length of "reasonably long" series of alike elements (successes) in a sequence of binary trials. More specifically, we are looking at the sum of the lengths of subsequences (strings) consisting of k or more consecutive successes (k is a given positive integer). It is clear that theoretical results on the distribution of this statistic are of major practical importance for establishing and investigating appropriate statistical tests which would detect the null hypothesis of randomness in the sequence of outcomes versus the alternative hypothesis of systematic clustering of successes. The investigation of the statistic of interest here is derived by exploiting a proper Markov chain embedding technique.

In Section 2 we present in brief the general family of Markov chain embeddable variables (c.f. Fu and Koutras (1994)) and introduce the wide subclass of Markov chain embeddable variables of polynomial type (MVP) which may be efficiently used for the evaluation of the exact distribution of enumerating random variables. This class generalizes the family of Markov chain embeddable variables of binomial type introduced in Koutras and Alexandrou (1995) and offers a very functional "working environment". Next we establish compact and computationally tractable formulae for obtaining the exact distribution (probability mass function and generating functions) of MVP's. In Section 3 we apply the MVP methodology to the problem of establishing the exact distribution of the sum of the lengths of success runs whose length exceed a prespecified level. Finally in Section 4 we assume that the composition of the observed sequence is known, that is to say, the number of successes and failures are fixed quantities, and proceed to the investigation of the conditional distribution of the aforementioned statistic. The results are then used (in Section 5) to investigate the performance of a new test of randomness.

2. General results

Recently, 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. 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 to the problem by exploiting proper extensions of the Markov chain embedding technique.

We shall first introduce the notion of a *Markov chain embeddable variable*, in a way similar to the one used by Fu and Koutras (1994). Let X_n (n a non-negative integer) be a non-negative finite integer-valued random variable and denote by $l_n = \sup\{x : \Pr(X_n = x) > 0\}$ its upper end point.

DEFINITION 2.1. The random variable X_n will be called a Markov Chain embeddable variable if

(a) there exists a Markov chain $\{Y_t : t \geq 0\}$ defined on a state space $\Omega = \{\alpha_1, \alpha_2, \dots\}$ which can be partitioned as $\Omega = \bigcup_{x \geq 0} C_x$,

(b) the probability mass function of X_n can be captured by considering the projection of the probability space of Y_n onto C_x, i.e.

$$\Pr(X_n = x) = \Pr(Y_n \in C_x), \qquad x = 0, 1, \dots, l_n.$$

If Λ_t is the one-step transition probability matrix of the Markov chain $(\{Y_t : t \geq 0\}, \Omega)$, then the exact distribution of X_n can be derived by the aid of the formula

$$\Pr(X_n = x) = \boldsymbol{\pi}_0 \left(\prod_{t=1}^{n} \Lambda_t \right) \sum_{i:\alpha_i \in C_x} \boldsymbol{e}_i', \qquad x = 0, 1, \dots, l_n \tag{2.1}$$

where $\boldsymbol{e}_i$ are unit (row) vectors and $\boldsymbol{\pi}_0 = (\Pr(Y_0 = \alpha_1), \Pr(Y_0 = \alpha_2), \Pr(Y_0 = \alpha_3), \dots)$ is the vector of initial probabilities.

For typographical convenience we impose the convention $\Pr(X_0 = 0) = 1$ and set $\prod_{t=a}^{b} \Lambda_t = I$ for $a > b$.

Should it be possible to partition Λ_t in the form of a bidiagonal blocked matrix with non-zero blocks appearing only on the main diagonal and the diagonal next to it, the investigation of X_n's distribution could be easily carried out by considering appropriate *probability vectors* describing the overall state formulation of the Markovian process at time t. Motivated by this observation, Koutras and Alexandrou (1995) proceeded to the introduction of a significant subclass of the family of Markov chain embeddable variables which offered a computationally efficient framework for tackling problems of this nature. Unfortunately, this class is not wide enough to accommodate the distributional problem we are aiming at in this article. For this reason we proceed to the introduction of a more general family of variables which will be called *Markov chain embeddable variables of polynomial type*.

To start with, let us first observe that without loss of generality we may assume that the state subspaces C_x, $x = 0, 1, \dots$ have the same finite cardinality $s = |C_x|$.

DEFINITION 2.2. The random variable X_n will be called Markov chain embeddable variable of polynomial type (MVP) if

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

$$\Omega = \bigcup_{x \geq 0} C_x, \qquad C_x = \{c_{x,0}, c_{x,1}, \ldots, c_{x,s-1}\}$$

(b) there exists a positive integer m such that for $t \geq 1$

$$\Pr(Y_t \in C_y \mid Y_{t-1} \in C_x) = 0 \quad \text{for all} \quad y \neq x, x+1, \ldots, x+m$$

(c) the probability mass function of X_n can be captured by considering the projection of the probability space of Y_n onto C_x i.e.

$$\Pr(X_n = x) = \Pr(Y_n \in C_x), \qquad n \geq 0,\ x \geq 0.$$

For $m = 1$, Definition 2.2 reduces to the definition of the Markov chain embeddable variables of binomial type introduced by Koutras and Alexandrou (1995).

Roughly speaking, a *MVP* is characterized by the following property: the state subclasses C_x, $x \geq 0$, can be ordered in such a way that once the chain enters C_x, the feasible one step transitions lead either to the same subclass C_x or to one of the next m (main state) subclasses $C_{x+1}, \ldots, C_{x+m}$. For the $m+1$ transition probability matrices

$$A_{t,i}(x) = (\Pr(Y_t = c_{x+i,j'} \mid Y_{t-1} = c_{x,j}))_{s \times s}, \qquad 0 \leq i \leq m,\ t \geq 1,\ x \geq 0$$

it is clear (c.f. condition (b) of Definition 2.2) that the matrix $\sum_{i=0}^{m} A_{t,i}(x)$ is stochastic. Moreover, on introducing the probability (row) vectors

$$\boldsymbol{f}_t(x) = (\Pr(Y_t = c_{x,0}), \Pr(Y_t = c_{x,1}), \ldots, \Pr(Y_t = c_{x,s-1})), \qquad t \geq 0,\ x \geq 0$$

it follows directly from condition (c) of Definition 2.2 that

$$\Pr(X_n = x) = \boldsymbol{f}_n(x)(1, 1, \ldots, 1)' = \boldsymbol{f}_n(x)\mathbf{1}', \qquad n \geq 0,\ x \geq 0.$$

Finally, convention $\Pr(X_0 = 0) = 1$ implies that

$$\boldsymbol{\pi}_0\mathbf{1}' = \boldsymbol{f}_0(0)\mathbf{1}' = (\Pr(Y_0 = c_{0,0}), \Pr(Y_0 = c_{0,1}), \ldots, \Pr(Y_0 = c_{0,s-1}))\mathbf{1}' = 1$$
$$\boldsymbol{\pi}_x\mathbf{1}' = \boldsymbol{f}_0(x)\mathbf{1}' = 0, \qquad x \geq 1.$$

Before proceeding to the development of general results facilitating the investigation of the exact distribution of a *MVP*, let us discuss in brief some potential applications where the approach taken here can be fruitfully used. Let $\mathcal{E}$ be an event (single or composite) associated with a sequence of binary (or multistate) trials and introduce a score function $f_i(\mathcal{E})$ denoting the points earned if event $\mathcal{E}$ occurs at the i-th trial. Then a *MVP* offers an appropriate methodological tool for investigating the exact distribution of the total score X_n achieved in a series of n outcomes. This setup is wide enough to accommodate the number of fixed length runs model ($f_i(\mathcal{E}) = 1$ if a run of prespecified length has been registered at the i-th trial), the sum of run lengths model introduced in Section 1 ($f_i(\mathcal{E})$ equals the exact length of a run completed at the i-th trial provided that the length exceeds a prespecified level), or even more complex models pertaining

to the occurrence of specific patterns. By way of example assume that we are trying to investigate the efficacy of n questions. Each time we observe a cluster of correct responses (e.g. a certain number of correct answers in a row, or segments containing a prespecified percentage of correct answers) c points are added to the subject's score, i.e. $f_i(\mathcal{E}) = c$. For each subsequent correct response d extra points are earned. Apparently, the total number of points collected upon the completion of the test may be considered as an indication of method's efficacy. The same statistic, under a slightly different description may be used for deciding whether a quality control process is out of control, whether a disease can be declared as contagious based on patterns of infected among non-infected plants in a transect through a field, etc. It is clear that in all these situations the knowledge of the distribution of the test statistic will help the practitioner to setup reasonable statistical procedures guarantying prespecified levels of type I error. The general results presented in this section are quite useful for the investigation of the aforementioned models and models of similar nature encountered in numerous areas of applied sciences.

Let us start our study with the next theorem which provides a method for the evaluation of the probability mass function of a *MVP*.

THEOREM 2.1. *The sequence of vectors* $\boldsymbol{f}_t(x)$ *satisfies the recurrence relation*

$$\boldsymbol{f}_t(x) = \sum_{i=0}^{\min(x,m)} \boldsymbol{f}_{t-1}(x-i)A_{t,i}(x-i), \qquad t \geq 1,\ x \geq 0.$$

PROOF. Let $t \geq 1$, $x \geq 0$ and $0 \leq j \leq s-1$. The total probability theorem yields

$$\Pr(Y_t = c_{x,j}) = \sum_{i=0}^{\min(x,m)} \sum_{r=0}^{s-1} \Pr(Y_t = c_{x,j} \mid Y_{t-1} = c_{x-i,r}) \Pr(Y_{t-1} = c_{x-i,r})$$

which can be equivalently written as

$$\begin{aligned} \Pr(Y_t = c_{x,j}) &= \sum_{i=0}^{\min(x,m)} \sum_{r=0}^{s-1} \boldsymbol{e}_{r+1} A_{t,i}(x-i)\boldsymbol{e}'_{j+1} \Pr(Y_{t-1} = c_{x-i,r}) \\ &= \sum_{i=0}^{\min(x,m)} \boldsymbol{f}_{t-1}(x-i)A_{t,i}(x-i)\boldsymbol{e}'_{j+1} \end{aligned}$$

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

Next, let $\varphi_t(z)$ and $\Phi(z,w)$ be the single and double generating functions

$$\varphi_t(z) = \sum_{x=0}^{\infty} \Pr(X_t = x)z^x, \qquad \Phi(z,w) = \sum_{t=0}^{\infty} \varphi_t(z)w^t$$

and denote by $\boldsymbol{\varphi}_t(z)$ and $\boldsymbol{\Phi}(z,w)$ the single (row) and double (row) vector generating functions of $\boldsymbol{f}_t(x)$, respectively, that is

$$\boldsymbol{\varphi}_t(z) = \sum_{x=0}^{\infty} \boldsymbol{f}_t(x)z^x, \quad t \geq 0, \qquad \boldsymbol{\Phi}(z,w) = \sum_{t=0}^{\infty} \boldsymbol{\varphi}_t(z)w^t.$$

It is clear that $\varphi_0(z) = \pi_0$, $\varphi_t(z) = \varphi_t(z)\mathbf{1}'$, $t \geq 1$ and $\Phi(z,w) = \Phi(z,w)\mathbf{1}'$.

We mention that, it is the rule rather than the exception that matrices $A_{t,i}(x)$ do not depend on x, that is $A_{t,i}(x) = A_{t,i}$ for all $t \geq 1$ and $x \geq 0$. In this case the vector generating function $\varphi_t(z)$ can be expressed in the form of a product as stated in the following theorem.

THEOREM 2.2. *If $A_{t,i}(x) = A_{t,i}$ for all $t \geq 1$ and $x \geq 0$ then the (single) vector generating function of X_t is given by*

$$\varphi_t(z) = \pi_0 \prod_{r=1}^{t} \left(\sum_{i=0}^{m} A_{r,i} z^i \right), \qquad t \geq 1.$$

PROOF. For $t \geq 1$ and upon using Theorem 2.1 we may write

$$\begin{aligned}
\varphi_t(z) &= \sum_{x=0}^{\infty} \boldsymbol{f}_t(x) z^x = \sum_{x=0}^{m} \sum_{i=0}^{x} \boldsymbol{f}_{t-1}(x-i) A_{t,i} z^x + \sum_{x=m+1}^{\infty} \sum_{i=0}^{m} \boldsymbol{f}_{t-1}(x-i) A_{t,i} z^x \\
&= \sum_{i=0}^{m} z^i \left(\sum_{x=i}^{m} \boldsymbol{f}_{t-1}(x-i) z^{x-i} \right) A_{t,i} + \sum_{i=0}^{m} z^i \left(\sum_{x=m+1}^{\infty} \boldsymbol{f}_{t-1}(x-i) z^{x-i} \right) A_{t,i} \\
&= \sum_{i=0}^{m} z^i \left(\sum_{y=0}^{\infty} \boldsymbol{f}_{t-1}(y) z^y \right) A_{t,i} = \varphi_{t-1}(z) \left(\sum_{i=0}^{m} A_{t,i} z^i \right).
\end{aligned}$$

The proof may easily be completed by repeated application of the last formula. □

It is well known that the probability generating function $\varphi(z)$ of the (generalized) binomial distribution is the product of the *binomial* terms $p_{r0} + zp_{r1}$, where p_{ri} denotes the probability of occurrence of outcome "i", $i = 0, 1$ at the r-th trial (two possible outcomes). Replacing the binomial terms by the *polynomial* terms $\sum_{i=0}^{m} p_{ri} z^i$, where p_{ri} denotes the probability of occurrence of outcome "i", $i = 0, 1, \ldots, m$ at the r-th trial ($m+1$ possible outcomes), we obtain the probability generating function of the *univariate multinomial* distribution introduced by Steyn (1956) (see also Panaretos and Xekalaki (1986) and Philippou *et al.* (1990)). In Theorem 2.2, $\sum_{i=0}^{m} p_{ri} z^i$ has been replaced by the polynomial term $\sum_{i=0}^{m} A_{r,i} z^i$, a fact justifying the nomenclature *polynomial type* used for the random variables studied here.

In the case of an homogeneous *MVP*, we have the next theorem.

THEOREM 2.3. *If $A_{t,i}(x) = A_i$ for all $t \geq 1$ and $x \geq 0$ then the double vector generating function of X_t can be expressed as*

$$\Phi(z,w) = \pi_0 \left(I - w \sum_{i=0}^{m} A_i z^i \right)^{-1}$$

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

PROOF. Follows readily from Theorem 2.2 on observing that

$$\Phi(z,w) = \sum_{t=0}^{\infty} \varphi_t(z) w^t = \pi_0 \sum_{t=0}^{\infty} \left(w \sum_{i=0}^{m} A_i z^i \right)^t = \pi_0 \left(I - w \sum_{i=0}^{m} A_i z^i \right)^{-1}$$

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

For an homogeneous *MVP* X_t let $\mu_t = E(X_t)$, $t \geq 1$, denote the mean of X_t and $M(w) = \sum_{t=1}^{\infty} \mu_t w^t$, its generating function. The next theorem provides two compact formulae for the evaluation of μ_t and $M(w)$ through the transition probability matrices A_i, $i = 0, 1, \ldots, m$.

THEOREM 2.4. *If* $A_{t,i}(x) = A_i$ *for all* $t \geq 1$ *and* $x \geq 0$ *then*

$$\mu_t = E(X_t) = \boldsymbol{\pi}_0 \left(\sum_{r=1}^{t} \left(\sum_{i=0}^{m} A_i \right) \right)^{r-1} \left(\sum_{i=1}^{m} i A_i \right) \mathbf{1}'$$

$$M(w) = \sum_{t=1}^{\infty} \mu_t w^t = \frac{w}{1-w} \boldsymbol{\pi}_0 \left(I - w \sum_{i=0}^{m} A_i \right)^{-1} \left(\sum_{i=1}^{m} i A_i \right) \mathbf{1}'.$$

PROOF. Exploiting the formula

$$\frac{d}{dz} \left(\sum_{i=0}^{m} A_i z^i \right)^t = \sum_{r=1}^{t} \left[\left(\sum_{i=0}^{m} A_i z^i \right)^{r-1} \left(\sum_{i=1}^{m} i A_i z^{i-1} \right) \left(\sum_{i=0}^{m} A_i z^i \right)^{t-r} \right]$$

we deduce, by virtue of Theorem 2.2

$$\mu_t = \frac{d}{dz} [\boldsymbol{\varphi}_t(z) \mathbf{1}'] \Big|_{z=1} = \boldsymbol{\pi}_0 \sum_{r=1}^{t} \left[\left(\sum_{i=0}^{m} A_i \right)^{r-1} \left(\sum_{i=1}^{m} i A_i \right) \left(\sum_{i=0}^{m} A_i \right)^{t-r} \right] \mathbf{1}'.$$

The first result follows readily by recalling that matrix $\sum_{i=0}^{m} A_i$ is stochastic.

The generating function $M(w)$ may be written as

$$\begin{aligned} M(w) &= \boldsymbol{\pi}_0 \sum_{t=1}^{\infty} \sum_{r=1}^{t} \left[\left(\sum_{i=0}^{m} A_i \right)^{r-1} \left(\sum_{i=1}^{m} i A_i \right) \right] w^t \mathbf{1}' \\ &= \boldsymbol{\pi}_0 w \sum_{r=1}^{\infty} \left(\sum_{i=0}^{m} A_i \right)^{r-1} w^{r-1} \sum_{t=r}^{\infty} w^{t-r} \left(\sum_{i=1}^{m} i A_i \right) \mathbf{1}' \end{aligned}$$

and the desired formula is effortlessly established by virtue of

$$\sum_{r=1}^{\infty} \left(\sum_{i=0}^{m} A_i \right)^{r-1} w^{r-1} = \left(I - w \sum_{i=0}^{m} A_i \right)^{-1}. \tag{2.2}$$

□

It goes without saying that for $m = 1$ the outcomes of Theorems 2.1–2.4 produce the respective results which have been developed by Koutras and Alexandrou (1997) for the Markov chain embeddable variables of binomial type.

3. The distribution of the sum of the exact lengths of runs of length at least k

Consider a sequence of Bernoulli trials $Z_1, Z_2, \ldots$ with success probabilities $p_t = \Pr(Z_t = 1)$, and failure probabilities $q_t = \Pr(Z_t = 0) = 1 - p_t$, $t \geq 1$ and let n, k be any positive integers with $n \geq k$. For $k \leq t \leq n$ we define

$$U_t = \begin{cases} k + \ell, & \text{if } \; Z_{t-k-\ell+1} = Z_{t-k-\ell+2} = \cdots = Z_t = 1 \text{ and } Z_{t-k-\ell} = Z_{t+1} = 0 \\ 0, & \text{otherwise} \end{cases}$$

(convention: $Z_0 = Z_{n+1} = 0$). Then the sum of the exact lengths of substrings of the sequence $Z_1, Z_2, \ldots, Z_n$ containing k or more consecutive successes, can be expressed as $X_n = \sum_{t=k}^{n} U_t$, $n \geq 1$. It is clear that the support of X_n is $\{0, k, k+1, \ldots, n\}$. For $n \leq k$ we set $X_n = 0$ and the support of X_n reduces to $\{0\}$.

In order to view the random variable X_n as a MVP, we set $C_x = \{c_{x,0}, c_{x,1}, \ldots, c_{x,k}\}$ where $c_{x,i} = (x, i)$, $0 \leq i \leq k$, $x \geq 0$ and define a Markov chain $\{Y_t, t \geq 0\}$ on $\Omega = \bigcup_{x \geq 0} C_x$ as follows: $Y_t = c_{x,i}$ (or equivalently $Y_t = (x, i)$) if in the first t outcomes, say $1001 \cdots 0\underbrace{11 \cdots 1}_{r}$, the observed sum of the exact lengths of runs of k or more consecutive successes is x and

$$i = \begin{cases} r, & \text{if } \; r = 0, 1, \ldots, k-1 \\ k, & \text{if } \; r \geq k. \end{cases}$$

It is apparent that, once the chain enters C_x, the one step transitions may lead only to the subclasses C_x, C_{x+1} or C_{x+k}. Hence the random variable X_n belongs to the class of MVP. The transition probability matrices $A_{t,i}$, $i = 0, 1, \ldots, k$, can be easily identified by observing that if $Y_t = c_{x,k}$ the feasible one step transitions of the chain lead either to substate $c_{x+1,k}$ (if $Z_{t+1} = 1$) or to substate $c_{x,0}$ (if $Z_{t+1} = 0$). Therefore, $A_{t,0}$ will be given by

$$A_{t,0} = \begin{bmatrix} q_t & p_t & 0 & \cdots & 0 & 0 & 0 \\ q_t & 0 & p_t & \cdots & 0 & 0 & 0 \\ \vdots & & & & & \vdots & \\ q_t & 0 & 0 & \cdots & 0 & p_t & 0 \\ q_t & 0 & 0 & \cdots & 0 & 0 & 0 \\ q_t & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}_{(k+1)\times(k+1)}$$

while $A_{t,2}, \ldots, A_{t,k-1}$ will be $(k+1) \times (k+1)$ matrices with all their entries 0. Matrix $A_{t,1}$ will have all its entries 0 except for the entry $(k+1, k+1)$ which equals p_t. Finally, $A_{t,k}$ will have all its entries 0 except for the entry $(k, k+1)$ which equals p_t. The appropriate initial probability vector of the Markov chain established here is given by $\boldsymbol{\pi}_0 = (1, 0, 0, \ldots, 0)$.

Recalling now Theorem 2.1 we may readily evaluate the probability mass function of X_n. Moreover exploiting Theorem 2.2 we may derive its probability generating function as

$$\varphi_n(z) = \sum_{x=0}^{\infty} P(X_n = x) z^x = \boldsymbol{\pi}_0 \prod_{r=1}^{n} (A_{r,0} + z A_{r,1} + z^k A_{r,k}) \mathbf{1}'.$$

In the case of iid trials with success probabilities p ($p_t = p$, $q_t = q$), Theorem 2.3 yields, after some routine calculations,

$$\Phi(z,w) = \sum_{n=0}^{\infty} \varphi_n(z) w^n = \boldsymbol{\pi}_0 (I - w(A_0 + zA_1 + z^k A_k))^{-1} \mathbf{1}' = \frac{P_1(z,w)}{P_2(z,w)} \tag{3.1}$$

where

$$\begin{aligned} P_1(z,w) &= 1 - wpz - (wp)^k (1 - z^k) - (wp)^{k+1} (z^k - z) \\ P_2(z,w) &= 1 - w(1 + pz) + w^2 pz + w^{k+1} qp^k (1 - z^k) + w^{k+2} qp^{k+1} (z^k - z). \end{aligned}$$

(I denotes the identity $(k+1) \times (k+1)$ matrix). It is not difficult to verify that $\Phi(z,w)$ may be written in the form

$$\Phi(z,w) = \sum_{n=0}^{\infty} \varphi_n(z) w^n = \frac{1 - wa_1(z) - w^k a_2(z) - w^{k+1} a_3(z)}{1 - [wb_1(z) + w^2 b_2(z) + w^{k+1} b_3(z) + w^{k+2} b_4(z)]}$$

where $a_i(z)$, $i = 1,2,3$ and $b_i(z)$, $i = 1,2,3,4$, are appropriate functions of z.

Following the methodology employed by Antzoulakos and Chadjiconstantinidis (2001), we may express $\varphi_n(z)$ as

$$\varphi_n(z) = \xi_0(z) - a_1(z)\xi_1(z) - a_2(z)\xi_k(z) - a_3(z)\xi_{k+1}(z)$$

where

$$\xi_i(z) = \sum_{n_1 + 2n_2 + (k+1)n_3 + (k+2)n_4 = n - i} \left(\sum_{j=1}^{4} n_j \right)! \prod_{j=1}^{4} \frac{b_j^{n_j}(z)}{n_j!}, \qquad i = 0, 1, k, k+1.$$

Since the generating function of $\varphi_n(z)$, $n \geq 0$, is a rational function of the form (3.1) a recursive scheme may be readily established by the aid of standard combinatorial techniques (see e.g. Chapter 4.1 in Stanley (1997)). More specifically we have the next result.

THEOREM 3.1. *If $Z_1, Z_2, \ldots, Z_n$ is a sequence of iid Bernoulli trials the probability generating function $\varphi_n(z)$ of the random variable X_n satisfies the recursive scheme*

$$\begin{aligned} \varphi_n(z) = {} & (1 + pz)\varphi_{n-1}(z) - pz\varphi_{n-2}(z) \\ & - qp^k (1 - z^k)\varphi_{n-k-1}(z) - qp^{k+1}(z^k - z)\varphi_{n-k-2}(z), \qquad n \geq k+2 \end{aligned}$$

with initial conditions

$$\varphi_n(z) = \begin{cases} 1, & \text{if } 0 \leq n < k \\ 1 - p^k + (pz)^k, & \text{if } n = k \\ 1 - p^k(1+q) + 2qp^k z^k + (pz)^{k+1}, & \text{if } n = k+1. \end{cases}$$

PROOF. The desired result follows by writing (3.1) in the form

$$P_2(z,w) \sum_{n=0}^{\infty} \varphi_n(z) w^n = P_1(z,w),$$

performing the multiplication in the LHS and considering the coefficients of w^n, $n = 0, 1, 2, \ldots$ in the resulting power series equality. □

As far as the probability mass function $g_n(x) = \Pr(X_n = x)$, $x \geq 0$ is concerned, its numerical computation can be easily achieved by launching the vector recursive scheme given in Theorem 2.1 (with matrices $A_{t,i}$, $i = 0, 1, \ldots, k$ being replaced by the special forms described earlier in this paragraph) and using the expression

$$\Pr(X_n = x) = \boldsymbol{f}_n(x)\mathbf{1}', \qquad n \geq 0,\ x = 0, 1, 2, \ldots, n;$$

note that, for $x = 1, 2, \ldots, k-1$ no calculations are necessary since in this range we always have $\boldsymbol{f}_n(x) = \mathbf{0}$.

In the special case of iid Bernoulli trials one could avoid working with vector recurrences. Instead he may exploit the following effective recursive scheme which ensues easily from the result established in Theorem 3.1.

THEOREM 3.2. *If $Z_1, Z_2, \ldots, Z_n$ is a sequence of iid Bernoulli trials, the probability mass function $g_n(x) = \Pr(X_n = x)$ of the random variable X_n satisfies the recursive scheme*

$$\begin{aligned} g_n(x) = {} & g_{n-1}(x) + pg_{n-1}(x-1) - pg_{n-2}(x-1) \\ & - qp^k(g_{n-k-1}(x) - g_{n-k-1}(x-k)) \\ & - qp^{k+1}(g_{n-k-2}(x-k) - g_{n-k-2}(x-1)), \qquad n \geq k+2,\ x \geq 0 \end{aligned}$$

with initial conditions

$$g_n(x) = 0, \quad \text{if } x < 0 \text{ or } x > n$$

$$g_n(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{if } x > 0 \end{cases} \quad \text{for } 0 \leq n < k$$

$$g_k(x) = \begin{cases} 1 - p^k, & \text{if } x = 0 \\ p^k, & \text{if } x = k \\ 0, & \text{if } 1 \leq x \leq k-1 \end{cases}$$

$$g_{k+1}(x) = \begin{cases} 1 - p^k(1+q), & \text{if } x = 0 \\ 2qp^k, & \text{if } x = k \\ p^{k+1}, & \text{if } x = k+1 \\ 0, & \text{if } 1 \leq x \leq k-1. \end{cases}$$

PROOF. It suffices to replace $\varphi_n(z)$, $n \geq 0$ in the recursive formula given in Theorem 3.1 by the power series

$$\varphi_n(z) = \sum_{x=0}^{\infty} g_n(x) z^x$$

and then consider the coefficients of z^x on both sides of the resulting identity. □

Although one can always resort to Theorem 2.2 to evaluate the exact distribution of X_n, when n and k become large the calculations might get time consuming. In this case

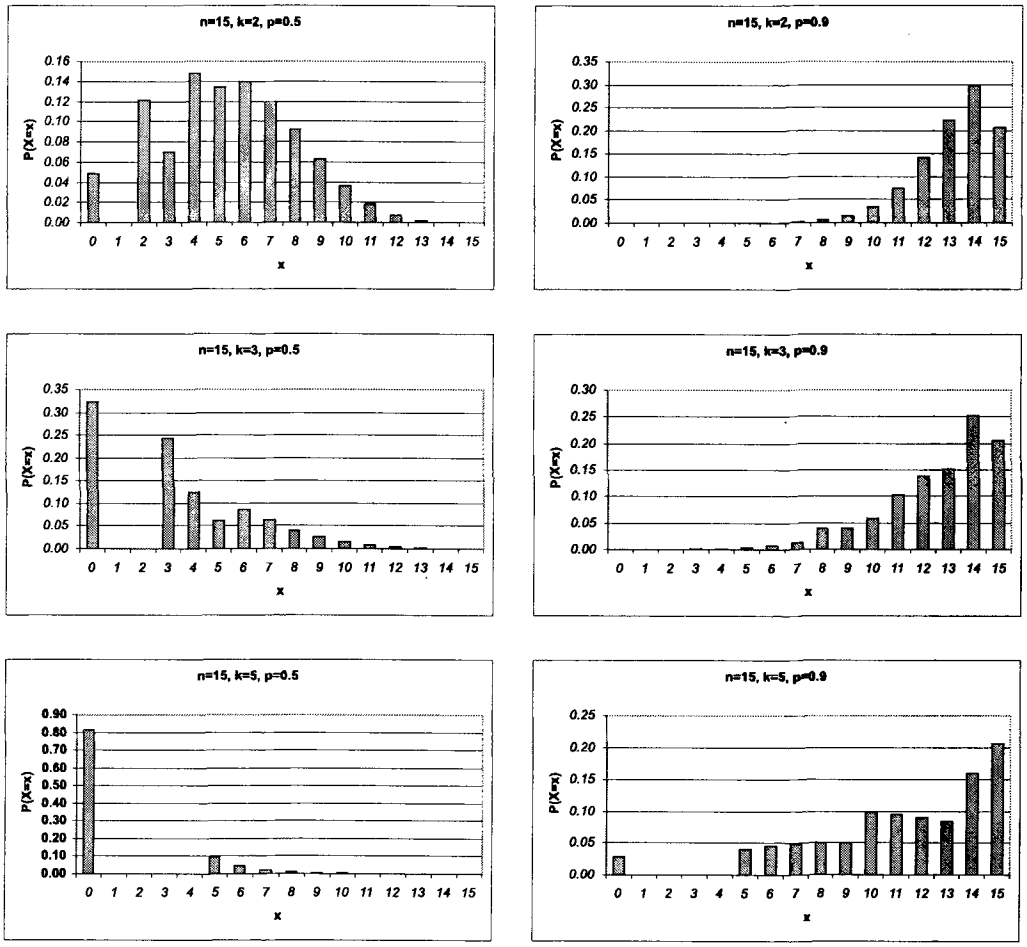

Fig. 1. Probability mass function of X_n for various n, k.

the investigation of the asymptotic distribution of X_n, will be quite valuable. Results of this flavour can be developed by appealing to the celebrated Chen-Stein method to settle an adequate Poisson convergence. Barbour *et al.* (1992) have provided a total variation distance bound for the joint distribution of runs of several lengths (see p. 244). Since applying a functional on the multivariate random variables involved does not increase the total variation distance, the upper bound offered there can be exploited for deriving an (asymptotic) estimate of the distribution of X_n. However we are not going to pursue this issue in the present article. The interested reader can urged to consult the monograph of Barbour *et al.* (1992) and work out the details of the aforementioned approach.

In Fig. 1 the probability mass function of X_n has been pictured for several values of n, k.

Theorem 3.1 can also be used for the derivation of a recursive formula for the raw moments of X_n. To this end, we observe first that the moment generating function $E[\exp(zX_n)]$ of X_n can be expressed as $E[\exp(zX_n)] = \varphi_n(e^z)$. Accordingly, replacing z by e^z in the recursive formula provided by Theorem 3.1 we may easily derive a recursive

scheme for it. Recalling next that

$$\mu_{n,r} = E(X_n^r) = \begin{cases} \dfrac{d^r}{dz^r} E[\exp(zX_n)]\Big|_{z=0}, & \text{if} \quad r \geq 1 \\ 1, & \text{if} \quad r = 0 \end{cases}$$

and making use of the well known formula

$$\frac{d^r}{dz^r}(e^{kz}E[\exp(zX_n)])\Big|_{z=0} = \sum_{i=0}^{r} \binom{r}{i} k^{r-i}\mu_{n,i}$$

we may readily verify the following theorem.

THEOREM 3.3. *The raw moments $\mu_{n,r}$, $r \geq 1$, of the random variable X_n satisfy the recursive scheme*

$$\mu_{n,r} = \mu_{n-1,r} + p\sum_{i=0}^{r} \binom{r}{i}(\mu_{n-1,i} - \mu_{n-2,i}) - qp^k\mu_{n-k-1,r}$$
$$+ qp^k \sum_{i=0}^{r} \binom{r}{i}(k^{r-i}(\mu_{n-k-1,i} - p\mu_{n-k-2,i}) + p\mu_{n-k-2,i}), \qquad n \geq k+2$$

with initial conditions

$$\mu_{n,r} = \begin{cases} 0, & \text{if} \quad 0 \leq n < k \\ k^r p^k, & \text{if} \quad n = k \\ 2k^r qp^k + (k+1)^r p^{k+1}, & \text{if} \quad n = k+1. \end{cases}$$

For $r = 1$, the aforementioned recursive scheme leads to the following second order difference equation for the sequence $\mu_s - \mu_{s-1} = E(X_s) - E(X_{s-1})$:

$$\mu_s - \mu_{s-1} = p(\mu_{s-1} - \mu_{s-2}) + qp^k(kq+p), \qquad s \geq k+2.$$

Applying this formula for $s = k+2, k+3, \ldots, n$ and summing up all the resulting equations we get

$$\mu_n - \mu_{k+1} = p(\mu_{n-1} - \mu_k) + (n-k-1)qp^k(kq+p), \qquad n \geq k+2.$$

If we replace next μ_k, μ_{k+1} (see Theorem 3.3) we may easily obtain the following first order difference equation for the means $\mu_n = E(X_n)$:

$$\mu_n = p\mu_{n-1} + kqp^k + (kq+p)p^k(p + (n-kq)), \qquad n \geq k+1.$$

For numerical calculation of μ_n, $n \geq 0$ by the aid of the last formula it suffices to recall the initial conditions

$$\mu_n = 0, \qquad 0 \leq n < k, \qquad \mu_k = kp^k.$$

Moreover, one could easily derive the solution of the above difference equation as

$$\mu_n = E(X_n) = p^k(k + (n-k)(kq+p)), \qquad n \geq k.$$

Needless to say, the same expression can be established by expanding the means generating function

$$M(w) = \sum_{n=0}^{\infty} E(X_n)w^n = \frac{(wp)^k(k - wp(k-1))}{(1-w)^2}$$

which is effortlessly deduced by a direct application of Theorem 2.4.

It is noteworthy that Theorem 3.1 can also be used for the derivation of a recursive scheme for the factorial moments

$$E(X_n(X_n-1)\cdots(X_n-r+1)) = \frac{d^r}{dz^r}\varphi_n(z)\bigg|_{z=1}.$$

The details are left to the reader.

Closing this section we mention that the approach used here for the study of the random variable X_n can be easily modified to cover the more general case where the sequence of trials exhibits a first order Markov dependence. To this goal, let us assume that $Z_1, Z_2, \ldots, Z_n$ is a sequence of Markov dependent trials, with transition probabilities defined by

$$p_{ij} = \Pr(Z_{t+1} = j \mid Z_t = i), \qquad t \geq 1,\ 0 \leq i,j \leq 1$$

and initial probabilities $\Pr(Z_1 = j) = p_j$, $j = 0,1$. Using exactly the same state definition as in the iid case it can be readily verified that the transition probability matrix $A_{t,0} = A_0$ takes on the form

$$A_0 = \begin{bmatrix} p_{00} & p_{01} & 0 & \cdots & 0 & 0 & 0 \\ p_{10} & 0 & p_{11} & \cdots & 0 & 0 & 0 \\ & \vdots & & & & \vdots & \\ p_{10} & 0 & 0 & \cdots & 0 & p_{11} & 0 \\ p_{10} & 0 & 0 & \cdots & 0 & 0 & 0 \\ p_{10} & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}_{(k+1)\times(k+1)}$$

while $A_{t,i} = A_t$, $i = 1,2,\ldots,k-1$, will be $(k+1)\times(k+1)$ matrices with all their entries 0. Matrix $A_{t,1} = A_1$ will have all its entries 0 except for the entry $(k+1, k+1)$ which equals p_{11}. Finally, $A_{t,k} = A_k$ will have all its entries 0 except for the entry $(k, k+1)$ which equals p_{11}.

Now making use of the formula

$$\Phi(z,w) = 1 + w\varphi_1(z)\left(I - w\sum_{i=0}^{m} A_i z^i\right)^{-1}\mathbf{1}'$$

with $\varphi_1(z) = (p_0, p_1, 0, \ldots, 0)_{1\times(k+1)}$ we get

$$\Phi(z,w) = \frac{Q_1(z,w)}{Q_2(z,w)}$$

where

$$Q_1(z;w) = 1 - w(\alpha + p_{11}z) + w^2\alpha p_{11}z - w^k p_1 p_{11}^{k-1}(1-z^k) - w^{k+1}p_{11}^k[p_1(z^k - z) + \gamma(1-z^k)] - w^{k+2}p_{11}^{k+1}\gamma(z^k - z),$$

$$Q_2(z;w) = 1 - w(1 + \alpha + p_{11}z) + w^2[\alpha + p_{11}z(1+\alpha)] - w^3\alpha p_{11}z + w^{k+1}\beta(1-z^k) + w^{k+2}\beta p_{11}(z^k - z)$$

and

$$\alpha = p_{11} - p_{01}, \qquad \beta = p_{10}p_{01}p_{11}^{k-1}, \qquad \gamma = p_{01} - p_1.$$

These expressions can be exploited to establish recurrence relations for the probability generating functions, probability mass functions and means of the random variable X_n under the Markovian set up. The interested reader may carry out the respective calculations in exactly the same way as in the iid case.

4. Conditional distribution

In this section it is assumed that the composition of the observed sequence is known, that is to say, the number of successes and failures are fixed quantities. The probabilities of our interest, therefore, become conditional ones. As elucidated in the next section outcomes of this nature are of primary interest in the development of tests of randomness in sequences of independent binary trials.

Let us assume again that we have a fixed number of Bernoulli trials $Z_1, Z_2, \ldots, Z_n$ with success probability $p = P(X_i = 1)$ and failure probability $q = P(X_i = 0) = 1 - p$, $i = 1, 2, \ldots, n$. Our intention is to investigate the conditional distribution of the run statistic X_n given the number $S_n = n - y$ $(0 \le y \le n)$ of successes in the n iid trials. Since S_n is a sufficient statistic for p, the conditional distribution we are looking at does not depend on p. In this section we shall use the notation $\Phi(z, w; p)$ and $\varphi_n(z; p)$ instead of $\Phi(z, w)$ and $\varphi_n(z)$, respectively, that is

$$\varphi_n(z;p) = \sum_{x=0}^{\infty} \Pr(X_n = x)z^x, \qquad \Phi(z,w;p) = \sum_{n=0}^{\infty} \varphi_n(z;p)w^n. \tag{4.1}$$

Let also

$$\psi_n(z;y) = \sum_{x=0}^{\infty} \Pr(X_n = x \mid S_n = n - y)z^x \tag{4.2}$$

denote the probability generating function of the conditional distribution of X_n given that $S_n = n - y$. The next theorem provides a formula for the double generating function of the quantity

$$a_n(z;y) = \binom{n}{y}\psi_n(z;y).$$

THEOREM 4.1. *The double generating function of* $a_n(z;y)$, $y = 0, 1, \ldots$, $n = y, y+1, \ldots$, *is given by*

$$\sum_{y=0}^{\infty}\left(\sum_{n=y}^{\infty} a_n(z;y)w^n\right)t^y = \Phi\left(z, (1+t)w; \frac{1}{1+t}\right).$$

PROOF. Replacing $\Pr(X_n = x)$ in (4.1) by the sum

$$\Pr(X_n = x) = \sum_{y=0}^{n} \Pr(X_n = x \mid S_n = n - y) \Pr(S_n = n - y)$$
$$= \sum_{y=0}^{n} \binom{n}{y} p^n \left(\frac{q}{p}\right)^y \Pr(X_n = x \mid S_n = n - y)$$

and making use of the expression (4.2) we deduce

$$\varphi_n(z; p) = \sum_{y=0}^{n} \binom{n}{y} p^n \left(\frac{q}{p}\right)^y \psi_n(z; y)$$

or equivalently

$$\Phi(z, w; p) = \sum_{n=0}^{\infty} \sum_{y=0}^{n} a_n(z; y) \left(\frac{q}{p}\right)^y (pw)^n.$$

Setting $t = q/p$ in the last expression we obtain

$$\Phi\left(z, w; \frac{1}{1+t}\right) = \sum_{n=0}^{\infty} \sum_{y=0}^{n} a_n(z; y) t^y \left(\frac{w}{1+t}\right)^n$$

and the required result follows immediately on replacing w by $(1 + t)w$. □

The outcome of Theorem 4.1 can be exploited to derive an explicit formula for the conditional distribution of X_n given the number of successes S_n. Specifically we have the following interesting result.

THEOREM 4.2. *The conditional probability* $\Pr(X_n = x \mid S_n = n - y)$ *is given by*

$$\Pr(X_n = x \mid S_n = n - y) = \binom{n}{y}^{-1} \sum_{r=0}^{y+1} \sum_{i=0}^{r} \sum_{j_1=0}^{r-i} \sum_{j_2=0}^{i} (-1)^{r+i+j_1-j_2}$$
$$\times \binom{y+1}{r} \binom{r}{i} \binom{r-i}{j_1} \binom{i}{j_2} \binom{r+a-1}{a} \binom{y+b}{b}$$

where $a = x - i + j_2 - k(j_1 + j_2)$ *and* $b = n - y - kr - i - a$.

PROOF. Making use of (3.1) we deduce

$$\Phi\left(z, (1+t)w; \frac{1}{1+t}\right) = \sum_{y=0}^{\infty} \left(\frac{1 - wz - w^k(1 - z^k) - w^{k+1}(z^k - z)}{(1 - w)(1 - wz)}\right)^{y+1} (wt)^y$$

which can be written in the form

$$\sum_{n=y}^{\infty} a_n(z; y) w^n = w^y \left(\frac{1 - wz - w^k(1 - z^k) - w^{k+1}(z^k - z)}{(1 - w)(1 - wz)}\right)^{y+1}$$
$$= w^y \left(\frac{1}{1 - w}\right)^{y+1} \left(1 - w^k \frac{(1 - z^k) + w(z^k - z)}{1 - wz}\right)^{y+1}.$$

Expanding the RHS in a power series with respect to w and employing the conventions $\binom{n}{m} = 0$ if $m < 0$ and $\binom{-1}{0} = 1$, we may easily arrive at the expression

$$a_n(z;y) = \sum_{m=0}^{\infty} \sum_{r=0}^{y+1} \sum_{i=0}^{r} (-1)^r \binom{y+s}{s} \binom{y+1}{r} \times \binom{r}{i} \binom{r+m-1}{m} (z^k - z)^i (1 - z^k)^{r-i} z^m, \qquad n \geq y$$

where $s = n - y - kr - i - m$.

The desired result follows immediately by a further expansion, by the aid of the binomial formula, of the powers appearing in the summand; c.f. (4.1), (4.2). □

5. Non-parametric tests of randomness

One of the widely known, oldest and easiest method of testing for random versus non-random ordering in a sequence of two types of symbols, is the classical *runs test* which has become a necessary addition in all contemporary non-parametric statistics textbooks, Bradley (1968), Gibbons and Chakraborti (1992). This test is based on the total number of runs, a run being any string of identical symbols which are followed and preceded by a different symbol or no symbol at all. An alternative test can be established by working with the length of the longest run in the observed sequence. Since an unusually long run indicates a tendency for like objects to cluster and, hence, the presence of a trend, Mosteller (1941) suggested a test for randomness based on the length of the longest run. The computation of the critical values of this test calls for the evaluation of the conditional distribution of the length of the longest run in n trials, given the number $S_n = n - y$ of the successes. We are now going to investigate a new test of randomness based on the statistic X_n introduced in the previous section (with k being a fixed pre-determined integer). Using an upper tailed test, the null distribution will be directly related to the conditional event

$$X_n \geq c \mid S_n = n - y,$$

where c is specified in terms of the significance level of the test. It is therefore apparent why an outcome like the one established in Theorem 4.2 is of special importance.

As mentioned earlier, in the 1940s, when the interest in the theory of runs was quite high, two different randomness tests were proposed: The classical runs test which was based on the total number R_n of runs of either type and the longest-run test which utilizes the length L_n of the longest success run. Recently, Agin and Godbole (1992) using the classical runs test as a model, developed a new exact test based on (a conditional version of) the total number $N_{n,k}$ of non-overlapping success runs of length k. This new test was found to be significantly more powerful in detecting certain types of clustering (non-randomness), than the classical runs test. Motivated by this result, Koutras and Alexandrou (1997), explored the performance of tests based on the total number $G_{n,k}$ of success runs of length at least k, and on the total number $M_{n,k}$ of overlapping success runs of length k. The test based on $M_{n,k}$ was found to be more powerful than the tests based on $G_{n,k}$ and $N_{n,k}$.

In the sequel, we conduct a systematic numerical experimentation in order to assess the performance of the randomness test based on the conditional distribution of the

Table 1. Empirical power/first-order Markov dependence model.

Parameters		$a = 0.10$		$a = 0.05$		$a = 0.01$	
p	n	X_n	$M_{n,k}$	X_n	$M_{n,k}$	X_n	$M_{n,k}$
0.99	50	0.98	0.99	0.99	1.00	0.96	0.88
0.99	100	0.98	0.98	0.98	0.97	0.98	0.85
0.99	150	0.94	1.00	0.97	1.00	0.93	0.81
0.95	50	0.92	0.92	0.91	0.92	0.90	0.76
0.95	100	0.86	0.99	0.80	0.95	0.87	0.87
0.95	150	0.82	0.98	0.74	0.97	0.69	0.91
0.65	50	0.74	0.59	0.46	0.25	0.39	0.14
0.65	100	0.62	0.65	0.48	0.54	0.39	0.14
0.65	150	0.63	0.66	0.49	0.54	0.41	0.15

Table 2. Empirical power/cyclical clustering model (with cycle length equal to 10).

Parameters		$a = 0.10$		$a = 0.05$		$a = 0.01$	
p	n	X_n	$M_{n,k}$	X_n	$M_{n,k}$	X_n	$M_{n,k}$
0.99	50	0.65	0.32	0.68	0.35	0.66	0.04
0.99	100	0.78	0.38	0.51	0.26	0.50	0.11
0.90	50	0.68	0.27	0.69	0.14	0.65	0.02
0.90	100	0.49	0.29	0.47	0.18	0.48	0.29
0.80	50	0.62	0.16	0.55	0.10	0.46	0.09
0.80	100	0.51	0.17	0.45	0.11	0.38	0.02
0.65	50	0.77	0.16	0.59	0.08	0.56	0.02
0.65	100	0.51	0.13	0.43	0.07	0.42	0.01

X_n (sum of the exact lengths of runs of length at least k). As already indicated, our tests are upper tailed and the critical values for rejection are determined by the aid of Theorem 4.2.

The empirical power of the new randomness test was compared to the empirical power of the randomness test based on $M_{n,k}$. The evaluation of the operational characteristics curves of the test, was achieved by the aid of Monte Carlo techniques. Thus, using specific alternatives 100 non random sequences were generated, and the probability (proportion) of rejecting the null hypothesis was computed.

The parametric configurations upon which the comparisons were performed are the following:

A. First-order Markov dependence: $p_1 = 0.5$ and

$$p_i = \begin{cases} p, & \text{if the } (i-1)\text{-th trial is a success} \\ p_1, & \text{if the } (i-1)\text{-th trial is a failure} \end{cases}$$

for $i = 2, 3, \ldots$.

B. Cyclical clustering (with cycle length equal to 10): The success probabilities

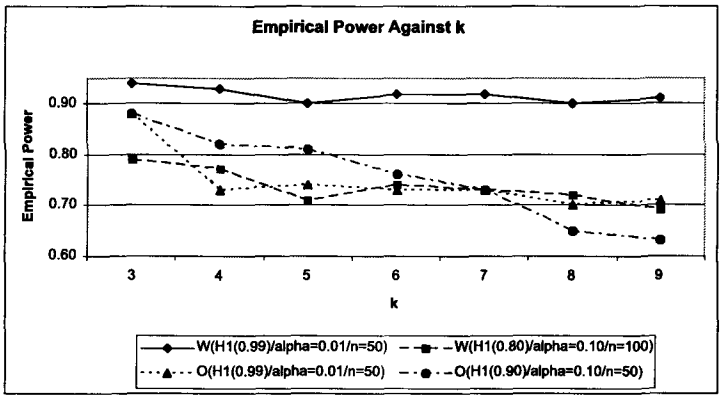

Fig. 2. Empirical power for various n, k.

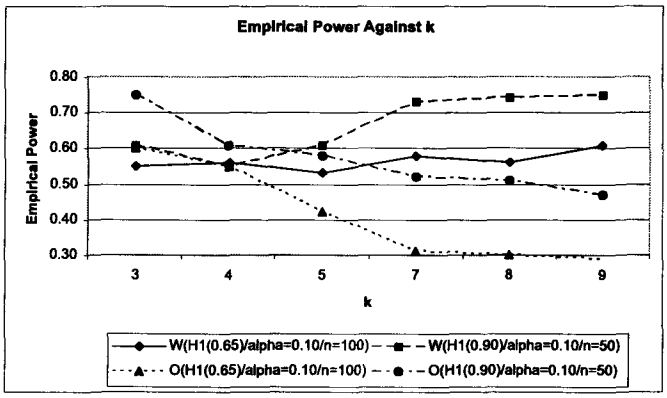

Fig. 3. Empirical power for various n, k.

p_i, $i = 1, 2, \ldots$ are given by

$$p_i = \begin{cases} p, & \text{if} \quad 10r + 1 \leq i \leq 10r + c, \ r = 0, 1, 2, 3, \ldots \\ 0.5, & \text{otherwise} \end{cases}$$

where $c \leq 10$ is a fixed integer.

It is worth mentioning that these sequences are indicative of real situations and have appeared before in certain practical applications (c.f. discussion in Koutras and Alexandrou (1997)).

The results of the simulation study, in the case where the sequence of outcomes exhibits a first-order Markov dependence, are displayed in Table 1. The empirical power recorded there was obtained by applying each test for all $k = 2, 3, \ldots, 9$ and choosing the largest power attained. Apparently the randomness tests based on X_n, are proved to be significantly more powerful than the ones based on $M_{n,k}$ when the type I error must be kept low ($a = 0.01$). For higher values of the significance level ($a = 0.05$ or $a = 0.10$) the performance of the two tests is comparable.

Shifting to the cyclical clustering model (Table 2) we observe that the X_n-based

Table 3. Empirical power against k/first-order Markov dependence model.

Parameters				k						
p	a	n	Statistic	2	3	5	6	7	9	max − min
0.90	.10	50	X_n	0.79	0.77	0.73	0.67	0.67	0.70	0.12
0.90	.10	50	$M_{n,k}$	0.82	0.82	0.74	0.78	0.61	0.62	0.21
0.90	.05	50	X_n	0.71	0.69	0.68	0.58	0.60	0.66	0.13
0.90	.05	50	$M_{n,k}$	0.86	0.82	0.67	0.63	0.48	0.54	0.38
0.95	.05	100	X_n	0.64	0.67	0.63	0.67	0.60	0.70	0.07
0.95	.05	100	$M_{n,k}$	0.88	0.82	0.71	0.65	0.59	0.50	0.38
0.65	.10	100	X_n	0.52	0.52	0.46	0.37	0.56	0.62	0.16
0.65	.10	100	$M_{n,k}$	0.65	0.60	0.41	0.26	0.32	0.28	0.39

test is always superior. In this case, there are instances where the $M_{n,k}$-based test leads to extremely low empirical values, while the new one attains significantly higher levels.

Closing we mention, that another interesting feature of the X_n-based test is that it is not very sensitive in changes on k, a property that is not present in the $M_{n,k}$-based test. This is clearly elucidated in Figs. 2 and 3 where the empirical powers (W stands for the X_n-based test and O for the $M_{n,k}$-based test) have been plotted against k (see also the results presented in Table 3). A direct consequence of this observation is that one has not to worry about the choice of k if he is going to use the X_n-based test. This is the main reason why we do not consider here the problem of developing empirical rules for the identification of reasonable values of k; such rules, when randomness tests based on fixed length runs are in use, have been given by Agin and Godbole (1992) and Koutras and Alexandrou (1997).

Acknowledgements

The authors wish to express their gratitude to the anonymous referees of this article. Based on their valuable and insightful comments, a substantial improvement of the original manuscript was made feasible.

References

Agin, M. A. and Godbole, A. P. (1992). A new exact runs test for randmness, *Computing Science and Statistcs* (eds. C. Page and R. Le Page), 281–285, Proceedings of the 22 Symposium on the Interface, Springer, New York.

Aki, S. and Hirano, K. (1988). Some characteristics of the binomial distribution of order k and related distributions, *Statistical Theory and Data Analysis II* (ed. K. Matusita), 211–222, Elsevier Science, North-Holland.

Antzoulakos, D. L. and Chadjiconstantinidis, S. (2001). Distributions of numbers of success runs of fixed length in Markov dependent trials, *Ann. Inst. Statist. Math.*, **53**, 599–619.

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

Balakrishnan, N., Balasubramanian, K. and Viveros, R. (1993). On sampling inspection plans based on the theory of runs, *Math. Sci.*, **18**, 113–126.

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

Balakrishnan, N., Mohanty, S. G. and Aki, S. (1997). Start-up demonstration tests under Markov dependence model with corrective actions, *Ann. Inst. Statist. Math.*, **49**, 155–169.

Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximations*, Oxford University Press, New York.

Bradley, J. (1968). *Distribution Free Statistical Tests*, Prentice Hall, New Jersey.

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.

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

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

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

Gibbons, J. D. and Chakraborti, S. (1992). *Nonparametric Statistical Inference*, 3rd ed., Marcel Deker, New York.

Godbole, A. P. (1990). On hypergeometric and related distributions of order k, *Comm. Statist. Theory Methods*, **19**, 1291–1301.

Godbole, A. P. (1992). The exact and asymptotic distribution of overlapping success runs, *Comm. Statist. Theory Methods*, **21**, 953–967.

Hahn, G. J. and Gage, J. B. (1983). Evaluation of a start-up demonstration test, *Journal of Quality Technology*, **15**, 103–105.

Han, Q. and Aki, S. (1999). Joint distributions of runs in a sequence of multistate trials, *Ann. Inst. Statist. Math.*, **51**, 419–447.

Hirano, K. and Aki, S. (1993). On 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.

Koutras, M. V. (1997). Waiting time distributions associated with runs of fixed length in two-state Markov chains, *Ann. Inst. Statist. Math.*, **49**, 123–139.

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. and Alexandrou, V. A. (1997). Nonparametric statistical randomness tests based on success runs of fixed length, *Statist. Probab. Lett.*, **32**, 393–404.

Mosteller, F. (1941). Note on an application of runs to quality control charts, *Ann. Math. Statist.*, **12**, 228–232.

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

Panaretos, J. and Xekalaki, E. (1986). On generalized binomial and multinomial distributions and their relation to generalized Poisson distributions, *Ann. Inst. Statist. Math.*, **38**, 223–231.

Philippou, A. N. and Makri, F. S. (1986). Success runs and longest runs, *Statist. Probab. Lett.*, **4**, 211–215.

Philippou, A. N., Antzoulakos, D. L. and Tripsiannis, G. A. (1990). Multivariate distributions of order k, Part II, *Statist. Probab. Lett.*, **10**, 29–35.

Stanley, R. P. (1997). *Enumerative Combinatorics*, Vol. I, Cambridge University Press.

Steyn, H. S. (1956). On the univariate series $F(t) = F(a; b_1, b_2, \ldots, b_k; t, t^2, \ldots, t^k)$ and its applications in probability theory, *Proc. Konink. Nederl. Akad. Wetensch. Ser. A*, **59**, 190–197.

Viveros, R. and Balakrishnan, N. (1993). Statistical inference from start-up demonstration test data, *Journal of Quality Technology*, **25**, 119–130.

Wolfowitz, J. (1943). On the theory of runs with some applications to quality control, *Ann. Math. Statist.*, **14**, 280–288.