Ann. Inst. Statist. Math.
Vol. 48, No. 4, 789–806 (1996)

ON A WAITING TIME DISTRIBUTION IN A SEQUENCE OF BERNOULLI TRIALS

M. V. KOUTRAS

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

(Received June 27, 1995; revised November 6, 1995)

Abstract. In the present article we investigate the exact distribution of the waiting time for the r-th non-overlapping appearance of a pair of successes separated by at most $k-2$ failures ($k \geq 2$) in a sequence of independent and identically distributed (iid) Bernoulli trials. Formulae are provided for the probability distribution function, probability generating function and moments and some asymptotic results are discussed. Expressions in terms of certain generalised Fibonacci numbers and polynomials are also included.

Key words and phrases: Waiting times, success runs, scans, distributions of order k, Fibonacci numbers, unimodality, Markov dependence.

1. Introduction

Let $X_1, X_2, \ldots$ be an infinite sequence of independent Bernoulli trials with constant success and failure probabilities $p = P(X_i = 1)$, $q = P(X_i = 0) = 1 - p$ respectively, $i = 1, 2, \ldots$. In the present paper we conduct a systematic study of the waiting time distribution for the r-th appearance of two successes which are separated by at most $k-2$ failures ($k \geq 2$, $r \geq 1$ are given integers). In coin tossing terminology, we may restate the problem as follows: If a p-biased coin is repeatedly tossed, and we place a marker each time we observe two heads in the last k (or less) successive flips, what is the probability to take n trials (tosses) until the placement of the r-th marker? It should be stressed that the enumerating scheme employed here is a non-overlapping one, in the sense that no success (head) contributes to more than one counts (markers). As an illustration, consider the sequence of outcomes $HTTHTHHTTHHHTHTTHHHH$ and $k = 3$. Then, the first 5 markers are placed at trials 6, 11, 14, 18 and 20 as shown below

$$HTT\overbrace{HTH}^{1}HTT\overbrace{HH}^{2}\overbrace{HTH}^{3}TT\overbrace{HH}^{4}\overbrace{HH}^{5}.$$

For $k = 2$, $r = 1$ we are in fact looking for the first success (head) run of length $k = 2$. The origins of this problem (in its more general form of success runs of length $k \geq 2$) should be attributed to De Moivre (see Todhunter (1965)).

Feller (1968) derived the probability generating function of the distribution as an application of the theory of recurrent events. For the last 15 years the term *geometric distribution of order k* (coined by Philippou *et al.* (1983)) has prevailed in the very extensive literature on this subject; more details can be found in Johnson *et al.* (1992).

For general $r \geq 1$ and $k = 2$ we have a special case of the *negative binomial distribution of order* $k = 2$ (see Philippou *et al.* (1983), Philippou (1984) and Hirano (1986)).

If $r = 1$, $k \geq 2$ the distribution under investigation is a special case of the detection waiting time when a 2-out-of-k *moving* (*sliding*) *window detector* is employed, Glaz (1983), Nelson (1978). Distributions of this type are of substantial interest in radar and safety systems, time-sharing computer networks etc. In reliability theory terminology, the respective tail probabilities are survival probabilities of *consecutive-2-within-k-out-of-n* reliability structures (see Chao *et al.* (1995), Papastavridis and Koutras (1993)). Also, the conditional tail distribution given the number of successes, is closely related to the *generalised birthday problem*, Naus (1968), Saperstein (1972). Additional applications pertaining to *quality control zone tests* can be found in Greenberg (1970), Roberts (1958) and Saperstein (1973).

The present paper is organised as follows. After the introduction of the necessary notations and definitions (Section 2), in Section 3 we conduct a systematic investigation of the waiting time distribution for the first occurrence ($r = 1$). Recursive relations are deduced for the moments and probability distribution function and its probability generating function is obtained. Moreover, the unimodality and the moment parameter estimation are discussed. The waiting time till the r-th occurrence is treated in Section 4, where an asymptotic (Poisson convergence) result is established as well. In Section 5 we examine the relation of the distribution to certain generalised Fibonacci-type numbers and polynomial. Finally, Section 6 furnishes some illustrative examples where the developed theory can be immediately applied whereas Section 7 states in brief a few results for the case of Markov dependent trials.

2. Definitions and notations

Let $X_1, X_2, \ldots$ be a sequence of independent Bernoulli trials with success and failure probabilities $p = P(X_i = 1)$, $q = P(X_i = 0) = 1 - p$ respectively and $k \geq 2$, $r \geq 1$ two positive integers. We shall denote by $T_{k,r}$ the waiting time for the r-th appearance of two successes which lie at most k places apart (separated by at most $k - 2$ failures). The probability distribution function and probability generating function of $T_{k,r}$ will be denoted by $f_{k,r}$ and $\phi_{k,r}$ respectively, i.e.

$$f_{k,r}(n) = P(T_{k,r} = n), \qquad n \geq 0$$

$$\phi_{k,r}(z) = E[z^{T_{k,r}}] = \sum_{n=0}^{\infty} f_{k,r}(n) z^n.$$

When no confusion is likely to arise we shall suppress the indices k, r using T, f, ϕ instead of $T_{k,1}$, $f_{k,1}$, $\phi_{k,1}$ and T_r, f_r, ϕ_r instead of $T_{k,r}$, $f_{k,r}$, $\phi_{k,r}$ respectively.

It is noteworthy that T can be formally defined as

$$T = \inf \left\{ n \geq 1 : \sum_{i=\max(1,n-k+1)}^{n} X_i \geq 2 \right\}.$$

For a fixed number of trials n, let $N_{n,k}$ denote the number of occurrences of a strand of k (at most) consecutive trials containing 2 successes, in the first n outcomes. Then the probability distributions of $T_{n,k}$ and $N_{n,k}$ are related by the obvious identity

$$P(N_{n,k} \geq r) = P(T_{k,r} \leq n). \tag{2.1}$$

Also, if L_n stands for the maximum number of successes appearing within any k consecutive trials in the first n outcomes, then

$$P(T > n) = P(N_{n,k} = 0) = P(L_n < 2).$$

3. Waiting time for the first occurrence

In this section the distribution of the random variable T is examined in some detail and several interesting properties of it are presented. Our first result provides an efficient recursive scheme for the evaluation of the probability distribution function $f(n)$.

THEOREM 3.1. *The probability distribution function $f(n) = P(T = n)$ of the waiting time random variable T satisfies the following recurrence relation*

$$f(n) = qf(n-1) + pq^{k-1} f(n-k), \qquad n > k \tag{3.1}$$

with initial conditions

$$\begin{aligned} &f(0) = f(1) = 0 \\ &f(n) = (n-1)p^2 q^{n-2}, \qquad 1 < n \leq k. \end{aligned} \tag{3.2}$$

PROOF. The derivation of initial conditions (3.2) is straightforward. Let us next assume that $n > k$. Manifestly

$$f(n) = P(T = n, X_1 = 0) + P(T = n, X_1 = 1)$$

and

$$P(T = n, X_1 = 0) = P(X_1 = 0)P(T = n \mid X_1 = 0) = qf(n-1).$$

On the other hand, the event $\{T = n, X_1 = 1\}$ is equivalent to $\{T = n\} \cap A$ with $A = \{X_1 = 1 \text{ and } X_i = 0 \text{ for all } 2 \leq i \leq k-1\}$ and therefore

$$P(T = n, X_1 = 1) = P(A)P(T = n \mid A)$$

which, on taking into account that $P(T = n \mid A) = f(n-k)$, yields

$$P(T = n, X_1 = 1) = pq^{k-1} f(n-k).$$

This completes the proof of formula (3.1).

For the special case $k = 2$, formula (3.1) reduces to

$$f(n) = qf(n-1) + pqf(n-2), \qquad n > 2$$

which is the well known recurrence for the geometric distribution of order $s = 2$ (see e.g. Shane (1973)). A by-product of the above is the alternative recursive relation

$$f(n) = f(n-1) - p^2 qf(n-3)$$

which is mentioned in Aki and Hirano (1989).

An immediate consequence of formula (3.1) is that the tail probabilities $\bar{F}(n) = P(T > n) = \sum_{x=n+1}^{\infty} f(x)$ satisfy exactly the same recurrence relation as $f(n)$. A direct proof of this appears in Roberts (1958) for $k = 3$ and Greenberg (1970) for general $k \geq 2$.

From (3.2) it is obvious that

$$\frac{f(n+1)}{f(n)} = \frac{n}{n-1} q, \qquad 1 < n \leq k-1 \tag{3.3}$$

and hence, for $n \leq k-1$ we have

$$\begin{aligned} f(n+1) &\geq f(n) \quad \text{for} \quad n \leq 1/p \\ f(n+1) &\leq f(n) \quad \text{for} \quad n \geq 1/p. \end{aligned}$$

On the other hand, (3.1), (3.2) guarantee that

$$f(n) \geq q^{k-1} f(n-k+1), \qquad n \geq k$$

and rewriting (3.1) as

$$f(n+1) - f(n) = p(-f(n) + q^{k-1} f(n-k+1)), \qquad n \geq k$$

we get

$$f(n+1) - f(n) \leq 0, \qquad n \geq k.$$

Accordingly, the distribution of T is unimodal, obtaining its maximum value for $n_0 = [\min(k-1, 1/p)] + 1$. Figure 1 shows the graphs of $f(n)$ for some typical values of k and p.

Using Theorem 3.1 we may also verify that $f(n)$ satisfies the strong unimodality characterization of Keilson and Gerber (1971)

$$f^2(n) \geq f(n-1) f(n+1)$$

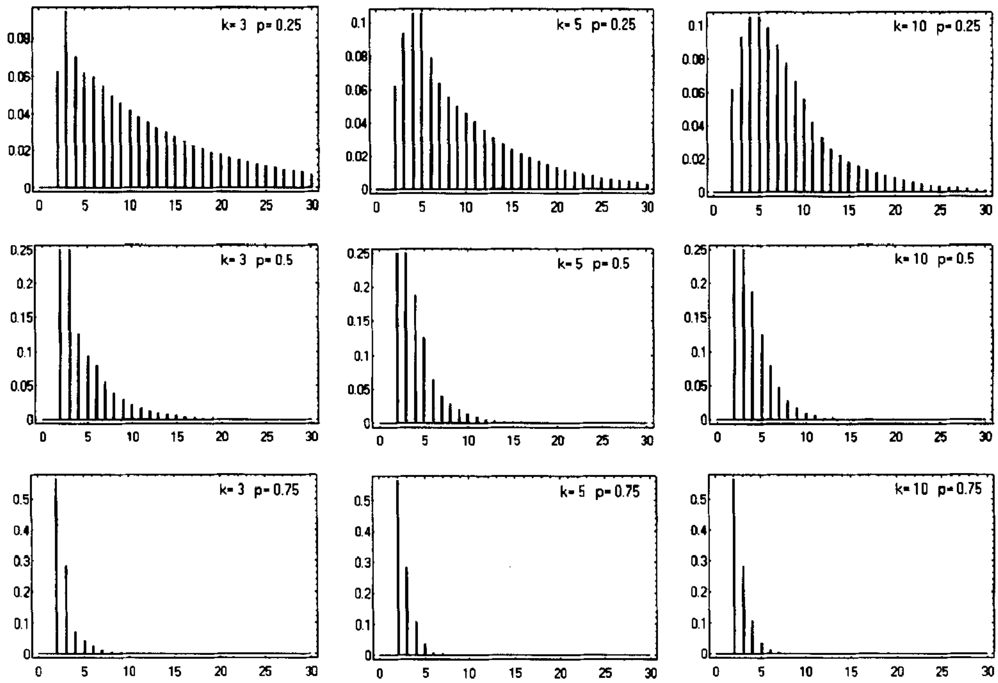

Fig. 1. Graphs of the probability distribution function $f(n) = P(T = n) = P(T_{k,1} = n)$.

for the range $2 \leq n \leq k$. However, for $n = k + 1$ the inequality is reversed, a fact implying that the distribution of T is not strongly unimodal; therefore its convolution with other unimodal distributions is not necessarily unimodal.

The probability generating function of T can be easily evaluated by manipulating over Theorem's 3.1 outcome. More specifically we have the next

THEOREM 3.2. *The probability generating function $\phi(z) = \sum_{n=0}^{\infty} f(n)z^n$ of T is given by*

$$\phi(z) = \frac{(pz)^2 A(z)}{1 - qz - pq^{k-1}z^k}, \qquad |z| \leq 1 \tag{3.4}$$

where

$$A(z) = \sum_{i=0}^{k-2} (qz)^i = \frac{1 - (qz)^{k-1}}{1 - qz}. \tag{3.5}$$

PROOF. Clearly, initial condition (3.2) is equivalent to

$$f(n) = qf(n-1) + p^2 q^{n-2}, \qquad 1 < n \leq k. \tag{3.6}$$

Writing $\phi(z)$ as

$$\phi(z) = \sum_{n=2}^{k} f(n)z^n + \sum_{n=k+1}^{\infty} f(n)z^n$$

and substituting $f(n)$ by (3.6) and (3.1) respectively, we get after some elementary algebra

$$\phi(z) = (pz)^2 A(z) + qz\phi(z) + pz(qz)^{k-1}\phi(z)$$

which proves the desired result.

An alternative approach for establishing formula (3.4) is to derive first the probability generating function $\bar{\phi}(z) = \sum_{n=0}^{\infty} \bar{F}(n)z^n$ of the tail probabilities $\bar{F}(n) = P(T > n)$ and subsequently employ the well known identity $\phi(z) = 1 + (z-1)\bar{\phi}(z)$. Two different methods for direct evaluation of $\bar{\phi}(z)$ can be found in Greenberg (1970) and Saperstein (1973).

The next corollary is an immediate by-product of Theorem 3.2.

COROLLARY 3.1. *The mean and variance of T are given by*

$$\mu = E[T] = \frac{2 - q^{k-1}}{p(1 - q^{k-1})}$$

$$\sigma^2 = \mathrm{Var}[T] = \frac{q}{p^2} + (2k-1)\frac{q^{k-1}}{p(1-q^{k-1})^2} + \frac{q}{p^2(1-q^{k-1})^2}.$$

PROOF. Both expressions result readily by evaluating the first two derivatives of $\phi(z)$ at $z = 1$ and substituting in the well known formulae

$$E[T] = \phi'(1), \qquad \mathrm{Var}[T] = \phi''(1) + \phi'(1) - (\phi'(1))^2.$$

The evaluation of higher order moments via the probability generating function (3.4) becomes rather cumbersome. However, it is not difficult to establish an easy to use recursive relation by manipulating directly on Theorem's 3.1 outcome. More precisely we have

THEOREM 3.3. *The moments of the random variable T about zero*

$$\mu'_s = E[T^s], \qquad s \geq 0$$

satisfy the next recurrence

$$\mu'_s = \frac{1}{p(1-q^{k-1})}\left\{p^2 \sum_{n=2}^{k} n^s q^{n-2} + q\sum_{i=0}^{s-1}\binom{s}{i}(1 + pq^{k-2}k^{s-i})\mu'_i\right\}, \qquad s \geq 1.$$

PROOF. Substituting (3.6) and (3.1) in the expression

$$\mu_s' = \sum_{n=2}^{k} n^s f(n) + \sum_{n=k+1}^{\infty} n^s f(n)$$

yields

$$\mu_s' = p^2 \sum_{n=2}^{k} n^s q^{n-2} + q \sum_{n=2}^{\infty} n^s f(n-1) + pq^{k-1} \sum_{n=k+1}^{\infty} n^s f(n-k). \tag{3.7}$$

The second and third sums, in view of the binomial formula, become

$$\begin{aligned} \sum_{n=1}^{\infty} (n+1)^s f(n) &= \sum_{i=0}^{s} \binom{s}{i} \mu_i' \\ \sum_{n=1}^{\infty} (n+k)^s f(n) &= \sum_{i=0}^{s} \binom{s}{i} k^{s-i} \mu_i' \end{aligned} \tag{3.8}$$

and the required result is easily built up by inplugging (3.8) in (3.7) and solving with respect to μ_s'.

It goes without saying that the results of Corollary 3.1 could alternatively be derived by applying Theorem 3.3 for $s = 1$ and $s = 2$. We also mention that for $k = 2$ Theorems 3.2 and 3.3 give the probability generating function and moment recurrence relations respectively of the geometric distribution of order 2.

Before closing the present section let us discuss briefly the problem of statistical estimation of the parameter p. Note first that the mean of T

$$E[T] = \frac{1}{p} \left[1 + \frac{1}{1 - (1-p)^k} \right] = h(p)$$

is a monotonically decreasing function in p, with

$$\lim_{p \to 0} h(p) = +\infty, \qquad \lim_{p \to 1} h(p) = 2.$$

If $T^{(1)}, T^{(2)}, \ldots, T^{(N)}$ is a random sample from this distribution, manifestly

$$\bar{T} = \frac{1}{N} \sum_{i=1}^{N} T^{(i)} \geq 2$$

and therefore the equation $h(p) = \bar{T}$ has a unique admissible root, which gives the moment estimator $\tilde{p}$ of p. Regarding the maximum likelihood estimation of p, since no simple analytic expression for $f(n)$ is available, the solution of the maximum likelihood equation has to be done iteratively. As mentioned by Aki and Hirano (1989) (in the parametric estimation problem for the class of binomial distributions of order k) the most important thing to this end is to develop efficient techniques for an easy and quick evaluation of the probability distribution function and its first and second derivatives (with respect to the parameter). Fortunately Theorem 3.1 offers an efficient scheme for computing $f(n)$. Yet, differentiating (3.1) and (3.2) two times yields equally simple recurrences permitting the fast numerical calculation of the first and second derivatives needed. The details are left to the reader.

4. Waiting for the r-th occurrence

We now turn our attention to the distribution of the waiting time $T_r = T_{k,r}$ for the r-th appearance of two successes separated by at most $k-2$ failures. We recall that we enumerate in a non-overlapping fashion, that is to say, the calculation procedure becomes anew each time a success gives birth to a count. With this in mind we may state that T_r can be decomposed in a sum of identical and independently distributed random variables with probability distribution function $f(n)$. To become more specific, if τ_1 denotes the waiting time for the first pair of successes lying k places apart, τ_2 the waiting time for the second one etc. it is evident that

$$T_r = \tau_1 + \tau_2 + \cdots + \tau_r \tag{4.1}$$

with τ_i being independent and following the distribution studied in the previous section.

THEOREM 4.1. *The probability generating function $\phi_r(z) = \sum_{n=0}^{\infty} f_r(n) z^n$ of T_r is given by*

$$\phi_r(z) = \left[\frac{(pz)^2 A(z)}{1 - qz - pq^{k-1}z^k}\right]^r, \qquad |z| \le 1 \tag{4.2}$$

where $A(z)$ is as in (3.5).

PROOF. The foregone discussion reveals that

$$\phi_r(z) = \prod_{i=1}^{r} E[z^{\tau_i}] = \prod_{i=1}^{r} E[z^T] = (\phi(z))^r$$

and (4.2) follows manifestly by employing (3.4).

The numerical evaluation of $f_r(n)$ can be easily achieved through the recurrences

$$f_{r+1}(n) - qf_{r+1}(n-1) - pq^{k-1} f_{r+1}(n-k) = p^2 \sum_{i=2}^{\min(n,k)} f_r(n-i) q^{i-2}$$

(convention: $f_r(n) = 0$ for $n < 0, r \geq 1$) and the initial conditions

$$f_r(n) = 0, \qquad 0 \leq n < 2r,\ r \geq 1$$

which are readily ascertainable from the obvious identity

$$(1 - qz - pq^{k-1}z^k)\phi_{r+1}(z) = (pz)^2 \left(\sum_{i=0}^{k-2} (qz)^i \right) \phi_r(z).$$

By virtue of (4.1) and Corollary 3.1, the mean and variance of T_r are

$$\begin{aligned} E[T_r] &= rE[T] = \frac{r(2 - q^{k-1})}{p(1 - q^{k-1})} \\ \operatorname{Var}[T_r] &= r\operatorname{Var}[T] = \frac{rq}{p^2} + (2k-1)\frac{rq^{k-1}}{p(1-q^{k-1})^2} + \frac{rq}{p^2(1-q^{k-1})^2}. \end{aligned}$$

It is perhaps unnecessary to point out that, as with the special case $r = 1$, there is a unique moment estimator of p whereas the numerical computations for its maximum likelihood estimator are highly facilitated from the aforementioned recurrences for $f_r(n)$.

For $k = 2$ the distribution of T_r reduces to the negative binomial distribution of order $k = 2$ which has been extensively studied recently after its introduction by Philippou *et al.* (1983). It is a well known fact that, if $rq \to \lambda > 0$ as $r \to \infty$, then this distribution (after being shifted to the support $0, 1, \ldots$) converges in law to the Poisson distribution of order $k = 2$; the last one is a special case of the class of generalised Poisson distributions with probability generating functions

$$\psi(z) = \psi(z; \lambda_1, \lambda_2, \ldots) = \exp\left(-\sum_{i=1}^{\infty} \lambda_i + \sum_{i=1}^{\infty} \lambda_i z^i \right), \qquad \sum_{i=1}^{\infty} \lambda_i < \infty$$

and probability distribution functions

$$P(Y = i) = \sum e^{-\sum_{j=1}^{\infty} \lambda_j} \frac{\prod_{j=1}^{i} \lambda_j^{y_j}}{\prod_{j=1}^{i} y_j!}$$

where the last summation is performed over all non-negative integers $y_1, y_2, \ldots, y_i$ such that $\sum_{s=1}^{i} s y_s = i$ (see Aki *et al.* (1984)).

The next theorem states that $T_{2,r}$ is the only random variable within the family $\{T_{k,r}, k \geq 2\}$ which converges in law to a pure ($\lambda_i > 0$ for at least one $i \geq 2$) generalised Poisson distribution.

THEOREM 4.2. *Assume that $rq \to \lambda > 0$ as $r \to \infty$. If $k = 2$ the random variable $T_{k,r} - 2r$ converges in law to the generalized Poisson distribution of order* 2 *with* $\lambda_1 = \lambda_2 = \lambda$, *otherwise* $(k > 2)$ *it is asymptotically distributed as an ordinary Poisson random variable with parameter* 2λ, *i.e.*

$$\lim_{r\to\infty} P(T_{k,r} - 2r = x) = \begin{cases} e^{-2\lambda}\dfrac{(2\lambda)^x}{x!} & \textit{for } k > 2 \\ e^{-2\lambda} \displaystyle\sum_{y_1+2y_2=x} \dfrac{\lambda^{y_1+y_2}}{y_1! y_2!} & \textit{for } k = 2. \end{cases}$$

Proof. The probability generating function of the random variable $Y = T_{k,r} - 2r$ is given by

$$\phi_Y(z) = z^{-2r}\phi_r(z) = \frac{p^{2r}A^r(z)}{(1 - qz - pq^{k-1}z^k)^r}.$$

But, under the assumptions made, we have

$$\lim_{r\to\infty} p^{2r} = \left(\lim_{r\to\infty}(1-q)^r\right)^2 = e^{-2\lambda}$$
$$\lim_{r\to\infty} r(qz + pq^{k-1}z^k) = \lambda(z + \delta_{k,2}z^k)$$
$$\lim_{r\to\infty} A^r(z) = \exp(-\lambda\delta_{k,2}z^{k-1} + \lambda z)$$

with δ_{ij} being Kronecker's delta (i.e. $\delta_{ii} = 1$ and $\delta_{ij} = 0$ for $i \neq j$). We can therefore write

$$\lim_{r\to\infty} \phi_Y(z) = \exp(-\lambda(1-z)(2 + \delta_{k,2}z^{k-1}))$$

or

$$\lim_{r\to\infty} \phi_Y(z) = \begin{cases} e^{-2\lambda(1-z)} = \psi(z; 2\lambda, 0, \ldots) & \text{for } k > 2 \\ e^{-\lambda(1-z)-\lambda(1-z^2)} = \psi(z; \lambda, \lambda, 0, \ldots) & \text{for } k = 2 \end{cases}$$

which establishes the required convergence.

It is well known (see e.g. Johnson *et al.* (1992)) that the shape factors

$$\beta_1 = \alpha_3^2, \qquad \beta_2 = \alpha_4$$

of a Poisson distribution obey the simple relation $\beta_2 - \beta_1 - 3 = 0$. For comparison reasons, in Table 1 we have listed the differences $\beta_2 - \beta_1$ between the shape factors of $T_{k,r}$ for a variety of r, p and k $(k > 2)$ values. The closeness of the (β_1, β_2) points to the "Poisson line" $\beta_2 - \beta_1 = 3$ is quite remarkable especially for large r values, as expected from Theorem's 4.2 output.

A further point of interest arising from Theorem 4.2 is that the asymptotic distribution of $T_{k,r}$, $k > 2$ is unimodal (as a matter of fact Poisson is a strongly unimodal distribution). On the contrary, for $k = 2$ neither $T_{2,r}$ nor its limit

Table 1. Differences $\beta_2 - \beta_1$ for the shape factors of $T_{k,r}$.

p	r	$k=3$	$k=4$	$k=5$	$k=10$	$k=20$	$k=100$
0.5	2	4.0222	4.0835	4.1645	3.8394	3.5033	3.5000
	5	3.4089	3.4334	3.4658	3.3358	3.2013	3.2000
	10	3.2044	3.2167	3.2329	3.1679	3.1007	3.1000
	100	3.0204	3.0217	3.0233	3.0168	3.0101	3.0100
	1000	3.0020	3.0022	3.0023	3.0017	3.0010	3.0010
0.8	2	4.4102	4.5329	4.1042	3.5024	3.5000	3.5000
	5	3.5641	3.6132	3.4417	3.2010	3.2000	3.2000
	10	3.2820	3.3066	3.2208	3.1005	3.1000	3.1000
	100	3.0282	3.0307	3.0221	3.0100	3.0100	3.0100
	1000	3.0028	3.0031	3.0022	3.0010	3.0010	3.0010
0.9	2	5.1141	4.3966	3.7304	3.5000	3.5000	3.5000
	5	3.8456	3.5586	3.2922	3.2000	3.2000	3.2000
	10	3.4228	3.2793	3.1461	3.1000	3.1000	3.1000
	100	3.0423	3.0279	3.0146	3.0100	3.0100	3.0100
	1000	3.0042	3.0028	3.0015	3.0010	3.0010	3.0010
0.95	2	5.9238	4.0640	3.5685	3.5000	3.5000	3.5000
	5	4.1695	3.4256	3.2274	3.2000	3.2000	3.2000
	10	3.5848	3.2128	3.1137	3.1000	3.1000	3.1000
	100	3.0585	3.0213	3.0114	3.0100	3.0100	3.0100
	1000	3.0058	3.0021	3.0011	3.0010	3.0010	3.0010
0.99	2	7.0889	3.6308	3.5031	3.5000	3.5000	3.5000
	5	4.6356	3.2523	3.2012	3.2000	3.2000	3.2000
	10	3.8178	3.1262	3.1006	3.1000	3.1000	3.1000
	100	3.0818	3.0126	3.0101	3.0100	3.0100	3.0100
	1000	3.0082	3.0013	3.0010	3.0010	3.0010	3.0010
0.999	2	7.4561	3.5135	3.5000	3.5000	3.5000	3.5000
	5	4.7824	3.2054	3.2000	3.2000	3.2000	3.2000
	10	3.8912	3.1027	3.1000	3.1000	3.1000	3.1000
	100	3.0891	3.0103	3.0100	3.0100	3.0100	3.0100
	1000	3.0089	3.0010	3.0010	3.0010	3.0010	3.0010

(Poisson distribution of order 2) are unimodal as the graphs in Hirano *et al.* (1984) reveal.

An additional asymptotic result could be stated by taking advantage of the fact that T_r is distributed as the sum of r independent variables. More specifically, in view of (4.1) and Corollary 3.1, the central limit theorem asserts that for fixed x we have

$$\lim_{r\to\infty} P(T_r < r\mu + x\sigma\sqrt{r}) = \Phi(x)$$

where $\Phi(x)$ denotes the cumulative distribution function of the standardised normal distribution.

Closing this section we give some results related to the number $N_{n,k}$ of occurrences of a strand of length (at most) k containing 2 successes in n Bernoulli trials.

THEOREM 4.3. *The double probability generating function of* $N_{n,k}$

$$\Phi(z,w) = \sum_{r=0}^{\infty} \sum_{n=2r}^{\infty} P(N_{n,k} = r) z^n w^r$$

is given by

$$\Phi(z,w) = \frac{1 + (p-q)z - pq^{k-1}z^k}{1 - 2qz + (q^2 - wp^2)z^2 - pq^{k-1}z^k + (q+wp)pq^{k-1}z^{k+1}}.$$

PROOF. Clearly

$$P(N_{n,k} = r) = P(N_{n,k} \geq r) - P(N_{n,k} \geq r+1)$$

and in view of (2.1) and Theorem 4.1 we may write

$$\sum_{n=2r}^{\infty} P(N_{n,k} = r) z^n = \phi^r(z) \frac{1 - \phi(z)}{1 - z}.$$

Accordingly

$$\Phi(z,w) = \frac{1 - \phi(z)}{1 - z} \cdot \frac{1}{1 - w\phi(z)}$$

and the required follows easily by substituting $\phi(z)$ and carrying out some elementary algebra.

COROLLARY 4.1. *The generating function of the means* $m_n = E[N_{n,k}]$ *is given by*

$$\sum_{n=0}^{\infty} m_n z^n = \left(\frac{pz}{1-z}\right)^2 \cdot \frac{1 - (qz)^{k-1}}{1 + (p-q)z - pq^{k-1}z^k}.$$

PROOF. It suffices to observe that

$$\sum_{n=0}^{\infty} m_n z^n = \left[\frac{\partial}{\partial w} \Phi(z,w)\right]_{w=1}$$

and make use of Theorem 4.3.

For the special case $k = 3$, the generating function formulae given in Theorem 4.3 and Corollary 4.1 coincide with the ones derived (by a different approach) in Koutras and Alexandrou (1995).

5. Fibonacci-type numbers and polynomials

For $k \geq 2$ a fixed positive integer let us introduce the sequence of numbers $\{F_{n,k}\}_{n\geq 0}$ or simply $\{F_n\}_{n\geq 0}$ by

$$\begin{aligned} &F_0 = F_1 = 0 \\ &F_n = F_{n-1} + 1, \qquad 2 \leq n \leq k \\ &F_n = F_{n-1} + F_{n-k}, \qquad n > k. \end{aligned} \tag{5.1}$$

Clearly $\{F_{n,2}\}_{n\geq 0}$ is a shifted version of the usual Fibonacci numbers; therefore, an appropriate name for the sequence $\{F_{n,k}\}_{n\geq 0}$ would be *k-step Fibonacci numbers*.

The next theorem expresses the distribution of $T_{k,r}$ in a *symmetric* sequence of Bernoulli trials (unbiased coin tossing) in terms of the k-step convoluted Fibonacci numbers.

THEOREM 5.1. *If* $p = q = 1/2$ *then*

$$f_r(n) = \frac{F_n^{(r)}}{2^n} \tag{5.2}$$

where $F_n^{(r)}$ *is the r-th convolution of the k-step Fibonacci numbers, i.e.*

$$F_n^{(i)} = \sum_{j=0}^{n} F_j^{(i-1)} F_{n-j}, \qquad i \geq 2. \tag{5.3}$$

(*Convention*: $F_n^{(1)} \equiv F_n$.)

PROOF. Applying Theorem 3.1 for $p = q = 1/2$ we can verify that the sequence $2^n f(n) = 2^n f_1(n)$ obeys exactly the same recurrences as $\{F_n\}_{n\geq 0}$. Accordingly

$$f(n) = \frac{F_n}{2^n}$$

and the generating function of $\{F_n\}_{n\geq 0}$ is easily shown to be (in lieu of Theorem 3.2)

$$\sum_{n=0}^{\infty} F_n z^n = \phi(2z) = \frac{z^2 \sum_{i=0}^{k-2} z^i}{1 - z - z^k}.$$

Now, the generating function of the r-th convoluted numbers $\{F_n^{(r)}\}_{n\geq 0}$ will be given by

$$\sum_{n=0}^{\infty} F_n^{(r)} z^n = \left(\sum_{n=0}^{\infty} F_n z^n \right)^r = \phi^r(2z)$$

which can be restated as

$$\sum_{n=0}^{\infty} \frac{F_n^{(r)} z^n}{2^n} = \phi^r(z). \tag{5.4}$$

The required identity (5.2) results immediately by a direct comparison of (5.4) to (4.2) specialised for $p = q = 1/2$.

So far we have touched only the probability distribution function of T for symmetric trials. Fortunately, the general case $(0 < p = 1 - q < 1)$ can be easily captured as well, by introducing sequences of polynomials instead of numbers. To this end, let us define the *k-step Fibonacci polynomials* $F_{n,k}(x) \equiv F_n(x)$ by

$$\begin{aligned} &F_0(x) = F_1(x) = 0 \\ &F_n(x) = F_{n-1}(x) + x^2, \qquad 2 \le n \le k \\ &F_n(x) = F_{n-1}(x) + xF_{n-k}(x), \qquad n > k. \end{aligned} \tag{5.5}$$

By the definition, the n-th k-step polynomial is of order $s+2$ where $sk+2 \le n \le sk+k+1$. As a consequence of recurrences (5.5) we get

$$\sum_{n=0}^{\infty} F_n(x) z^n = \frac{(zx)^2 \sum_{i=0}^{k-2} z^i}{1 - z - xz^k} = F(z,x)$$

and the next theorem can be easily verified by following an exact parallel to that of Theorem 5.1.

THEOREM 5.2. *If $F_n^{(r)}(x)$ is the r-th convolution of the k-step polynomials, i.e.*

$$F_n^{(i)}(x) = \sum_{j=0}^{n} F_j^{(i-1)}(x) F_{n-j}(x), \qquad i \ge 2$$

then the probability distribution function of T_r can be expressed as

$$f_r(n) = P(T_r = n) = q^n F_n^{(r)}(p/q).$$

(*Convention*: $F_n^{(1)}(x) \equiv F_n(x)$.)

PROOF. Note first that

$$\sum_{n=0}^{\infty} F_n^{(r)}(x) z^n = \left(\sum_{n=0}^{\infty} F_n(x) z^n \right)^r = F^r(z,x)$$

and make use of the fact that $F^r(qz, p/q)$ coincides with $\phi_r(z)$ of Theorem 4.1.

6. Applications

When dealing with experimental trials with two possible outcomes (success/failure, good/bad, acceptable/not acceptable), it is of great significance to be able to develop criteria providing evidence of clustering of one of the two types of outcomes. Perhaps the oldest and most commonly used procedures of this type are the ones based on the concept of success runs. The random variable $T_{k,r}$

studied in the present paper is offering an efficient alternative test statistic in a variety of situations where the classical run criteria have been in use. A few specific examples related to applied work will now be discussed.

A method to study spatial patterns (diffusion of species, spread of diseases etc.) in ecology is to sample over belt transects and analyse the registered sequence of outcomes. Assume that we wish to study whether a certain disease is able to spread up to a k-step distance. Certainly, the observation of two infected trees lying k steps apart is not providing enough evidence for our suspicion; however, should r such pairs be observed, it is sensible to stop sampling and reject the null hypothesis if the waiting time $T_{k,r}$ is reasonably large. Setting up the critical region as $T_{k,r} > c$, we can easily determine c so that a specific significance level is achieved, by making use of the theory developed in Sections 4 and 5. The aforementioned technique provides an alternative to the run-based procedures used in testing segregation between two species (see Pielou (1963, 1977)).

In the same spirit, $T_{k,r}$ could be used in educational psychology studies of transfer and learning to decide whether a particular treatment should be terminated or not. For details on related run and scan-based techniques on this subject we refer to Bogartz (1965), Glaz (1989, 1993), Fu and Koutras (1994), Koutras and Alexandrou (1995) and Koutras *et al.* (1994, 1995).

Another interesting application is offered by the following simple variant of the moving (sliding) window detection problem appearing in Nelson (1978) and Glaz (1983). Consider a radar sweep with a quantizer transmitting to the detector the digit 1 or 0 according to whether the signal-plus-noise waveform exceeds a predetermined threshold. The detector's memory keeps track of the last k (at most) transmitted digits and generates a pulse when two 1's are observed. Should this happen, the contents of detector's memory are erased and the next transmitted digit is the first to be registered. The occurrence of the r-th pulse initiates an alarm. Manifestly, the results of the previous sections provide means for studying the waiting time for an alarm (in the case of iid transmissions).

A final application comes from the area of statistical quality control. Greenberg (1970) and Saperstein (1973) examined zone tests in which a process is declared "out of control" if in a subsequence of k consecutive sampled items there exist at least $s \leq k$ observations outside the zone (say, the three sigma limits about the mean). The material of the present paper (Sections 4 and 5) is closely related to a modified sampling plan which seeks an assignable cause whenever r disjoined pairs of defective items are spotted, with the elements of each pair being at most k places apart (i.e. separated by $k-2$ good items).

Closing we mention that $T_{k,r}$ can be easily associated to certain start-up demonstration tests analogous to the ones introduced and studied by Balakrishnan *et al.* (1995) and Viveros and Balakrishnan (1993). Results on this topic will be reported in detail in a forthcoming paper.

7. A Markov dependent model

Recently, quite a few research work has been done on run-related problems in sequences of trials where the outcome of a trial depends on the outcomes of the previous trials in a Markovian fashion, Aki and Hirano (1993), Balasubramanian *et al.* (1993), Hirano and Aki (1993), Uchida and Aki (1995). The methodology employed for the derivation of the results of Sections 3–4 is easily amenable for tackling the more general waiting time problem which arises by dropping the independence assumption and replacing the Bernoulli trials by a sequence of first order Markov dependent outcomes. In the sequel we are stating in brief some formulae which hold true under this set-up.

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

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

and initial probabilities $P(X_0 = j) = \delta_{0,j}$, $j = 0, 1$. Retaining the notations of Section 2 for the waiting time random variable and the respective probability distribution (generating) function of the Markovian model we may state the following:

1. The probability distribution function $f(n) = P(T = n)$ satisfies the recurrence relation

$$f(n) = p_{00} f(n-1) + p_{01} p_{10} p_{00}^{k-2} f(n-k), \qquad n > k$$

with initial conditions

$$\begin{aligned} &f(0) = f(1) = 0 \\ &f(n) = (n-2) p_{01}^2 p_{10} p_{00}^{n-3} + p_{01} p_{11} p_{00}^{n-2}, \qquad 2 \leq n \leq k. \end{aligned}$$

2. The probability generating function $\phi(z) = \sum_{n=0}^{\infty} f(n) z^n$ of T is given by

$$\phi(z) = \frac{p_{01} z [p_{11} z + p_{10} p_{01} z^2 \sum_{i=3}^{k} (p_{00} z)^{i-3}]}{1 - p_{00} z - p_{01} p_{10} p_{00}^{k-2} z^k}, \qquad |z| \leq 1.$$

3. The probability generating function $\phi_r(z) = \sum_{n=0}^{\infty} f_r(n) z^n$ of T_r is given by

$$\phi_r(z) = \phi^r(z), \qquad |z| \leq 1.$$

4. If $rp_{00} \to \lambda > 0$ and $rp_{10} \to \mu > 0$ as $r \to \infty$ then for the probability generating function $\phi_Y(z)$ of the shifted random variable $Y = T_{k,r} - 2r$ we have

$$\lim_{r \to \infty} \phi_Y(z) = \exp(-(1-z)[(\lambda + \mu) + \mu \delta_{k,2} z^{k-1}])$$

or

$$\lim_{r \to \infty} \phi_Y(z) = \begin{cases} e^{-(\lambda+\mu)(1-z)} = \psi(z; \lambda + \mu, 0, \ldots) & \text{for } k > 2 \\ e^{-\lambda(1-z) - \mu(1-z^2)} = \psi(z; \lambda, \mu, 0, \ldots) & \text{for } k = 2. \end{cases}$$

Accordingly, for $k = 2$ the random variable $T_{k,r} - 2r$ converges in law to the generalized Poisson distribution of order 2 with $\lambda_1 = \lambda$, $\lambda_2 = \mu$, whereas for $k > 2$ it is asymptotically distributed as an ordinary Poisson random variable with parameter $\lambda + \mu$.

The proofs of (i)–(iv) are easily verified by employing exactly the same arguments as the ones used for the case of independent Bernoulli trials (see Sections 3–4); the technical details are left to the reader.

Acknowledgements

The exposition of the paper has been greatly benefited from referees' critical comments and suggestions.

The present 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. and Hirano, K. (1989). Estimation of parameters in the discrete-distributions of order k, *Ann. Inst. Statist. Math.*, **41**, 47–61.

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.

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

Balasubramanian, K., Viveros, R. and Balakrishnan, N. (1993). Sooner and later waiting time problems for Markovian Bernoulli trials, *Statist. Probab. Lett.*, **18**, 153–161.

Bogartz, R. (1965). The criterion method: some analyses and remarks, *Psychological Bulletin*, **64**, 1–14.

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

Glaz, J. (1983). Moving window detection for discrete-data, *IEEE Trans. Inform. Theory*, **29**, 457–462.

Glaz, J. (1989). Approximations and bounds for the distribution of the scan statistic, *J. Amer. Statist. Assoc.*, **84**, 560–566.

Glaz, J. (1993). Approximations for the tail probabilities and moments of the scan statistic, *Statistics in Medicine*, **12**, 1845–1852.

Greenberg, I. (1970). The first occurrence of n successes in N trials, *Technometrics*, **12**, 627–634.

Hirano, K. (1986). Some properties of the distributions of order k, *Fibonacci Numbers and their Applications* (eds. A. N. Philippou, A. F. Horadam and G. E. Bergum), 43–53, Reidel.

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

Hirano, K., Kuboki, H., Aki, S. and Kuribayashi, A. (1984). Figures of probability density functions in statistics II—discrete univariate case, *Computer Science Monographs*, No. 20, The Institute of Statistical Mathematics, Tokyo.

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

Keilson, J. and Gerber, H. (1971). Some results on discrete unimodality, *J. Amer. Statist. Assoc.*, **66**, 386–389.

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 Academic Publishers, Netherlands.

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

Naus, J. (1968). An extention of the birthday problem, *Amer. Statist.*, **22**, 27–29.

Nelson, J. B. (1978). Minimal-order models for false-alarm calculations on sliding windows, *IEEE Trans. Aerospace Electron. Systems*, **14**, 351–363.

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, Amsterdam.

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.

Pielou, E. C. (1963). Runs of healthy and diseased trees in transects through an infected forest, *Biometrics*, **19**, 603–614.

Pielou, E. C. (1977). *Mathematical Ecology*, Wiley, New York.

Roberts, S. W. (1958). Properties of control zone tests, *Bell Systems Technical Journal*, **37**, 83–114.

Saperstein, B. (1972). The generalized birthday problem, *J. Amer. Statist. Assoc.*, **67**, 425–428.

Saperstein, B. (1973). On the occurrence of n successes within N Bernoulli Trials, *Technometrics*, **15**, 809–818.

Shane, H. D. (1973). A Fibonacci probability function, *Fibonacci Quart.*, **11**, 517–522.

Todhunter, I. (1965). *A History of the Mathematical Theory of Probability*, Chelsea, New York.

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.

Viveros, R. and Balakrishnan, N. (1993). Statistical inference from start-up demonstration test data, *Journal of Quality Technology*, **22**, 119–130.