J. Appl. Prob. **40,** 346–360 (2003)
Printed in Israel

ON EXACT AND LARGE DEVIATION APPROXIMATION FOR THE DISTRIBUTION OF THE LONGEST RUN IN A SEQUENCE OF TWO-STATE MARKOV DEPENDENT TRIALS

JAMES C. FU * ** AND
LIQUN WANG,* *University of Manitoba*
W. Y. WENDY LOU,*** *University of Toronto*

Abstract

Consider a sequence of outcomes from Markov dependent two-state (success–failure) trials. In this paper, the exact distributions are derived for three longest-run statistics: the longest failure run, longest success run, and the maximum of the two. The method of finite Markov chain imbedding is used to obtain these exact distributions, and their bounds and large deviation approximation are also studied. Numerical comparisons among the exact distributions, bounds, and approximations are provided to illustrate the theoretical results. With some modifications, we show that the results can be easily extended to Markov dependent multistate trials.

Keywords: Success run; failure run; Markov chain imbedding; transition probability matrix; probability of large deviation

AMS 2000 Subject Classification: Primary 60E05
Secondary 60J10

1. Introduction

Let random variables $\{X_t\}_{t=1}^n$ be outcomes of a sequence of random trials with two states $\mathcal{F} = \{\mathrm{F}, \mathrm{S}\}$, which are either independent and identically distributed (i.i.d.) two-state trials or homogeneous Markov dependent trials with transition probability matrix

$$\boldsymbol{A} = \begin{bmatrix} p_{\mathrm{FF}} & p_{\mathrm{FS}} \\ p_{\mathrm{SF}} & p_{\mathrm{SS}} \end{bmatrix}. \tag{1.1}$$

Given $n \in J^+ = \{1, 2, \dots\}$, we define

$$L_n(\mathrm{S}) = \max_{1 \le t \le n-k+1} \{k : X_t = X_{t+1} = \cdots = X_{t+k-1} = \mathrm{S}\}$$

and

$$L_n(\mathrm{F}) = \max_{1 \le t \le n-k+1} \{k : X_t = X_{t+1} = \cdots = X_{t+k-1} = \mathrm{F}\}$$

as the lengths of the longest success and failure runs respectively. Further, we define

$$L_n = \max\{L_n(\mathrm{S}), L_n(\mathrm{F})\}$$

as the length of the longest run, of either successes or failures.

Received 30 September 2002; revision received 8 January 2003.
* Postal address: Department of Statistics, University of Manitoba, Winnipeg, Manitoba R3T 2N2, Canada.
** Email address: fu@cc.umanitoba.ca
*** Postal address: Department of Public Health Sciences, University of Toronto, Toronto, Ontario M5S 1A8, Canada.

The concept of the longest run has long been studied, due to its mathematical challenge and extensive applications in, for example, game theory, biology, reliability theory, and quality control. There are very many results on the exact distribution of the length of the longest success run, especially when $\{X_t\}$ is a sequence of i.i.d. two-state trials.

Let $A_n(k)$ be the number of sequences of length n in which the longest success run does not exceed k. Then, for the probability of success $p = \frac{1}{2}$, $A_n(k)$ satisfies the recursive equation

$$A_n(k) = \begin{cases} \sum_{j=0}^{k} A_{n-1-j}(k) & \text{for } k < n, \\ 2^n & \text{for } k = n \end{cases}$$

(see Schilling (1990)) and the distribution of $L_n(\mathrm{S})$ is given by

$$\mathrm{P}(L_n(\mathrm{S}) \le k) = 2^{-n} A_n(k). \tag{1.2}$$

Recently, Suman (1994) related the probability of the longest run of either successes or failures in n Bernoulli trials to the longest success in $n-1$ Bernoulli trials with $p = \frac{1}{2}$ as

$$\mathrm{P}(L_n \le k) = \mathrm{P}(L_{n-1}(\mathrm{S}) \le k-1) \tag{1.3}$$

for $k = 1, \ldots, n$. In this special case, the probability $\mathrm{P}(L_n \le k)$ can be computed via (1.2). Note, however, that the relationship in (1.3) holds only in the case when the $\{X_t\}$ are from Bernoulli trials with $p = \frac{1}{2}$.

For general cases $0 < p < 1$, Burr and Cane (1961) and Gibbons (1971) showed that the distribution of $L_n(\mathrm{S})$ can be written as

$$\mathrm{P}(L_n(\mathrm{S}) \le k) = \sum_{i=0}^{n} C_n^{(i)}(k) p^i q^{n-i},$$

where $q = 1 - p$ and

$$C_n^{(i)}(k) = \sum_{j=0}^{i} C_{n-1-j}^{(i-j)}(k).$$

Philippou and Makri (1986) and Hirano (1986) independently provided the exact formula

$$\mathrm{P}(L_n(\mathrm{S}) < k) = \sum_{i=0}^{n} p^i q^{n-i} \sum_{j=0}^{k} \sum_{x_1,\ldots,x_{k+1}} \binom{x_1 + \cdots + x_{k+1}}{x_1, \ldots, x_{k+1}}$$

for $k = 0, 1, \ldots, n$, where the inner summation is defined over all nonnegative integers $x_1, \ldots, x_{k+1}$ satisfying the conditions that $x_1 + 2x_2 + \cdots + (k+1)x_{k+1} = n - j$ and $x_1 + \cdots + x_{k+1} = n - i$, and where

$$\binom{x_1 + \cdots + x_{k+1}}{x_1, \ldots, x_{k+1}} = \frac{(x_1 + \cdots + x_{k+1})!}{\prod_{i=1}^{k+1} (x_i)!}$$

is the multinomial coefficient. These combinatoric formulae are rather complex and can be very tedious in computation.

For large n, there are several outstanding results on the length of the longest success run. For example, Petrov (1965), Rényi (1970), and Erdős and Révész (1975) showed that, as $n \to \infty$,

$$\frac{L_n(\mathrm{S})}{\log_{1/p} n} \longrightarrow 1 \quad \text{almost surely.} \tag{1.4}$$

The result of (1.4) is often referred to as the new law of large numbers. Goncharov (1944) proved that $L_n(\mathrm{S}) - \log_{1/p} n$ does not have a proper limiting distribution, and showed that the sequence of probabilities follows the law:

$$\mathrm{P}(L_n(\mathrm{S}) - \lfloor \log_{1/p} n \rfloor < x) = \exp\{-n(1-p)p^{\lfloor \log_{1/p} n \rfloor + x}\} + o(1), \tag{1.5}$$

where $\lfloor a \rfloor$ stands for the integer part of a and $x = 0, \pm 1, \pm 2, \ldots$. In their recent book, Balakrishnan and Koutras (2002, pp. 29–35) provided up-to-date results and information on the asymptotic distribution of $L_n(\mathrm{S})$ in the i.i.d. cases.

In recent studies of the reliability of consecutive-k-out-of-n:F systems, Fu (1986), Chao and Fu (1989), (1991), and Papastavridis and Koutras (1993) provided uniform upper and lower bounds for the tail probability of $L_n(\mathrm{S}) < k$:

$$(1-p^k)^{n-k+1} \le \mathrm{P}(L_n(\mathrm{S}) < k) \le (1-qp^k)^{n-k+1} \tag{1.6}$$

for $k = 1, \ldots, n$ and $0 < p < 1$. The bounds are very simple and accurate, especially when p is small or q is close to 1. It is worth mentioning that the upper bound of (1.6), $(1-qp^k)^{n-k+1}$, is equivalent to the approximation (1.5) given by Goncharov (1944), whereas the lower bound $(1-p^k)^{n-k+1}$ with $p = (\lambda t/n)^{1/k}$ converges to the tail probability of an exponential distribution. The details of these interesting connections will be explained in Sections 4 and 5.

Most results in the literature have focused on the longest success run. Recently, the exact distribution of the longest run in a sequence of multistate trials has been studied by Fu (1996), using the finite Markov chain imbedding technique. Using the same technique, Lou (1996) studied the joint distribution of the longest success run $L_n(\mathrm{S})$ and the number of successes S_n, and also the conditional distribution of $L_n(\mathrm{S})$ given S_n for the sequence $\{X_i\}$ under i.i.d. (Bernoulli) or Markov dependent bi-state trials. To date, the exact distribution of L_n under Markov dependent trials remains unknown, and when n (with k fixed) is large, no good approximation for $\mathrm{P}(L_n < k)$ has been developed. The purpose of this work is to develop a simple and efficient method for finding the exact distributions of $L_n(\mathrm{S})$, $L_n(\mathrm{F})$, and L_n for small and moderate n. In the case where n is large, we develop a large deviation approximation which will provide a good approximation for $\mathrm{P}(L_n < k)$ and will also capture its exponential rate tending to zero as $n \to \infty$.

The combinatoric method is, in general, inefficient and cumbersome in studying the exact distributions of the random variables $L_n(\mathrm{F})$, $L_n(\mathrm{S})$, and L_n, especially when the sequence $\{X_t\}$ is Markov dependent. In Section 2, we derive the exact distributions of the three longest-run statistics through the adoption of the finite Markov chain imbedding technique developed and used by Fu (1986), (1996), Fu and Koutras (1994), Koutras (1997), Koutras and Alexandrou (1997), Doi and Yamamoto (1998), and Boutsikas and Koutras (2000). Existing bounds for $L_n(\mathrm{S})$ are reviewed in Section 3. In Section 4, we use the finite Markov chain imbedding technique combined with spectral analysis to obtain large deviation approximations for the tail probabilities of the longest runs. In Section 5, some numerical results are provided to illustrate our theories. Our numerical results show that the upper bound of (1.6) and the large

deviation approximation perform very well. Possible extension of our results to multistate Markov dependent trials is discussed in Section 6.

2. Exact distributions

Given $k \in J^+ = \{1, 2, \ldots\}$, let $\Lambda_0(k) = \mathrm{F}^k = \mathrm{FF}\ldots\mathrm{F}$ and $\Lambda_1(k) = \mathrm{S}^k = \mathrm{SS}\ldots\mathrm{S}$ denote two simple patterns of size k, and let $\Lambda(k) = \Lambda_0(k) \cup \Lambda_1(k)$ be the compound pattern generated by $\Lambda_0(k)$ and $\Lambda_1(k)$ (i.e. either $\Lambda_0(k)$ or $\Lambda_1(k)$ occurs). We define three waiting time random variables corresponding to the three patterns $\Lambda_0(k)$, $\Lambda_1(k)$, and $\Lambda(k)$, respectively, as

$$W_0(k) = \inf\{n : n \in J^+, X_{n-k+1} = \cdots = X_n = \mathrm{F}\},$$

$$W_1(k) = \inf\{n : n \in J^+, X_{n-k+1} = \cdots = X_n = \mathrm{S}\},$$

$$W(k) = \inf\{n : n \in J^+, X_{n-k+1} = \cdots = X_n = \mathrm{F} \text{ or } X_{n-k+1} = \cdots = X_n = \mathrm{S}\}.$$

Since $\Lambda(k)$ is a compound pattern generated by $\Lambda_0(k)$ and $\Lambda_1(k)$, it follows from the definitions of $W_0(k)$, $W_1(k)$, and $W(k)$ that

$$W(k) = \min\{W_0(k), W_1(k)\}.$$

It is known (Fu and Chang (2002)) that waiting time random variables of simple or compound patterns are finite Markov chain imbeddable in the following sense: for any waiting time random variable $W(\Lambda)$ defined by a pattern $\Lambda = \bigcup_{i=1}^{l} \Lambda_i$ (simple or compound), there exists a Markov chain $\{Y_t\}$ defined on a finite state space $\Omega = \{1, 2, \ldots, m, \alpha_1, \alpha_2, \ldots, \alpha_l\}$ with transition probability matrix $\boldsymbol{M}$, which has the form

$$\boldsymbol{M} = \left[\begin{array}{c|c} \boldsymbol{N} & \boldsymbol{C} \\ \hline \mathbf{0} & \boldsymbol{I} \end{array}\right]_{(m+l)\times(m+l)},$$

where $\mathbf{0}$ is the null matrix, $\mathbf{I}$ is the identity matrix,

$$\boldsymbol{N} = \begin{bmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{bmatrix}, \quad \text{and} \quad \boldsymbol{C} = \begin{bmatrix} p_{1\alpha_1} & \cdots & p_{1\alpha_l} \\ \vdots & \ddots & \vdots \\ p_{m\alpha_1} & \cdots & p_{m\alpha_l} \end{bmatrix}.$$

States $\alpha_1, \ldots, \alpha_l$ are absorbing states corresponding to patterns $\Lambda_1, \ldots, \Lambda_l$ respectively. The waiting time distribution of $W(\Lambda)$ is then given by

$$\mathrm{P}(W(\Lambda) = n) = \boldsymbol{\xi}_0 \boldsymbol{N}^{n-1}(\boldsymbol{I} - \boldsymbol{N})\mathbf{1}^\top, \qquad n = 1, 2, \ldots, \tag{2.1}$$

where $\boldsymbol{\xi}_0$ is the initial distribution and $\mathbf{1}^\top$ is the transpose of the row vector $\mathbf{1} = [1, 1, \ldots, 1]_{1\times m}$. The sequence $\{Y_t\}$ is referred to as the imbedded Markov chain corresponding to the waiting time random variable $W(\Lambda)$. In our context, there are three imbedded Markov chains, $\{Y_t(0)\}$, $\{Y_t(1)\}$, and $\{Y_t\}$, corresponding to the waiting time random variables $W_0(k)$, $W_1(k)$, and $W(k)$ respectively, defined on three state spaces:

$$\Omega_0 = \{\phi, \mathrm{S}, \mathrm{F}\} \cup \mathcal{F}(\Lambda_0) = \{\phi, \mathrm{S}, \mathrm{F}, \mathrm{F}^2, \ldots, \mathrm{F}^{k-1}, \alpha_0\},$$

$$\Omega_1 = \{\phi, \mathrm{S}, \mathrm{F}\} \cup \mathcal{F}(\Lambda_1) = \{\phi, \mathrm{F}, \mathrm{S}, \mathrm{S}^2, \ldots, \mathrm{S}^{k-1}, \alpha_1\},$$

$$\Omega = \Omega_0 \cup \Omega_1 = \{\phi, \mathrm{F}, \mathrm{F}^2, \ldots, \mathrm{F}^{k-1}, \mathrm{S}, \mathrm{S}^2, \ldots, \mathrm{S}^{k-1}, \alpha_0, \alpha_1\},$$

where ϕ is the initial state (dummy state), $\mathcal{F}(\Lambda_i)$ stands for the set of all sequential subpatterns of pattern Λ_i, and F^i and S^i are, respectively, the failure run and success run of length i. The transition probability matrix of the imbedded Markov chain $\{Y_t(0)\}$ is given by

$$\boldsymbol{M}_0 = \left[\begin{array}{c|c} \boldsymbol{N}_0 & \boldsymbol{C}_0 \\ \hline \mathbf{0} & \boldsymbol{I} \end{array}\right], \tag{2.2}$$

where $\mathbf{0}$ is a row vector of zeros, $\boldsymbol{I}$ is the (1×1) identity matrix,

$$\boldsymbol{N}_0 = \begin{bmatrix} 0 & p_{\mathrm{S}} & p_{\mathrm{F}} & 0 & 0 & \cdots & 0 \\ 0 & p_{\mathrm{SS}} & p_{\mathrm{SF}} & 0 & 0 & \cdots & 0 \\ 0 & p_{\mathrm{FS}} & 0 & p_{\mathrm{FF}} & 0 & \cdots & 0 \\ 0 & p_{\mathrm{FS}} & 0 & 0 & p_{\mathrm{FF}} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & p_{\mathrm{FS}} & 0 & 0 & \vdots & \ddots & p_{\mathrm{FF}} \\ 0 & p_{\mathrm{FS}} & 0 & 0 & 0 & \cdots & 0 \end{bmatrix},$$

and $\boldsymbol{C}_0 = (0, 0, \ldots, 0, p_{\mathrm{FF}})^\top$. The transition probability matrices $\boldsymbol{M}_1$ and $\boldsymbol{N}_1$ of $\{Y_t(1)\}$ have the same forms as $\boldsymbol{M}_0$ and $\boldsymbol{N}_0$, but with S and F interchanged. Similarly, the transition probability matrix of $\{Y_t\}$ is

$$\boldsymbol{M} = \left[\begin{array}{c|c} \boldsymbol{N} & \boldsymbol{C} \\ \hline \mathbf{0} & \boldsymbol{I} \end{array}\right], \tag{2.3}$$

where

$$\boldsymbol{N} = \begin{bmatrix} 0 & p_{\mathrm{F}} & 0 & \cdots & 0 & p_{\mathrm{S}} & 0 & \cdots & 0 \\ 0 & 0 & p_{\mathrm{FF}} & \ddots & \vdots & p_{\mathrm{FS}} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & p_{\mathrm{FF}} & p_{\mathrm{FS}} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & p_{\mathrm{FS}} & 0 & \cdots & 0 \\ 0 & p_{\mathrm{SF}} & 0 & \cdots & 0 & 0 & p_{\mathrm{SS}} & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & 0 & \ddots & 0 \\ 0 & p_{\mathrm{SF}} & 0 & \cdots & 0 & 0 & \vdots & \ddots & p_{\mathrm{SS}} \\ 0 & p_{\mathrm{SF}} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

and

$$\boldsymbol{C}^\top = \begin{bmatrix} 0 & \cdots & 0 & p_{\mathrm{FF}} & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & p_{\mathrm{SS}} \end{bmatrix}.$$

It can be seen that the transition probability matrix $\boldsymbol{M}$ is a combination of the two matrices $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$. With this notation, we have the following result.

Theorem 2.1. *Suppose that $\{X_t\}_{t=1}^n$ is a Markov chain with transition probability matrix $\boldsymbol{A}$ given by (1.1). Then the exact distribution of the longest run L_n is given by*

$$\mathrm{P}(L_n < k) = \boldsymbol{\xi}_0 \boldsymbol{N}^n \mathbf{1}^\top \tag{2.4}$$

for $k = 2, \ldots, n$, where the transition matrix $\boldsymbol{N}$ is defined by (2.3).

Proof. This is an immediate consequence of the equality

$$\mathrm{P}(L_n < k) = \mathrm{P}(W(k) > n)$$

and (2.1).

Note that, when $k = 1$, $\mathrm{P}(L_n < 1) = 0$.

Similarly, the exact distributions for $L_n(\mathrm{F})$ and $L_n(\mathrm{S})$ are given by

$$\mathrm{P}(L_n(\mathrm{F}) < k) = \boldsymbol{\xi}_0 \boldsymbol{N}_0^n \mathbf{1}^\top \tag{2.5}$$

and

$$\mathrm{P}(L_n(\mathrm{S}) < k) = \boldsymbol{\xi}_0 \boldsymbol{N}_1^n \mathbf{1}^\top \tag{2.6}$$

for $k = 1, \ldots, n$, where $\boldsymbol{N}_0$ and $\boldsymbol{N}_1$ are defined by (2.2).

Given p_{FF}, p_{FS}, p_{SF}, and p_{SS}, the equations (2.4), (2.5), and (2.6) are very efficient formulae for evaluating the exact distributions numerically. Since the formulae involve only the power of the transition probability matrix of the imbedded Markov chain, this clearly paves the way for spectral analysis and large deviation approximations, as will be seen in Section 4.

3. Upper and lower bounds for the longest success and failure runs

It can be shown that the probability that the longest failure run $L_n(\mathrm{F})$ is less than k is equal to the reliability of consecutive-k-out-of-n:F systems (Fu (1986) and Fu and Koutras (1994)):

$$R(k, n) = \mathrm{P}(L_n(\mathrm{F}) < k).$$

In reliability literature, various upper and lower bounds for $R(k, n)$ have been obtained for the case where the components are i.i.d. (i.e. Bernoulli trials). Basically there are three types of bounds:

(i) Bonferroni-type inequalities given by Derman *et al.* (1982);

(ii) product-type inequalities

$$(1 - q^k)^{n-k+1} \le \mathrm{P}(L_n(\mathrm{F}) < k) \le (1 - pq^k)^{n-k+1} \tag{3.1}$$

for $k = 1, \ldots, n$ and $0 < p < 1$, studied by Fu (1986), Chao and Fu (1989), and Papastavridis and Koutras (1993);

(iii) Stein–Chen-type inequalities,

$$|\mathrm{P}(L_n(\mathrm{F}) < k) - \exp(-\lambda_n)| \le (2kp - 1)q^k \tag{3.2}$$

for small q ($q \sim c/n^{1/k}$), where $\lambda_n = (n-k+1)pq^k$, studied by Barbour, Chryssaphinou and Roos (1995), Chryssaphinou and Papastavridis (1990), Barbour, Holst and Janson (1992), and Gordon *et al.* (1986).

Similar to (3.1), we have

$$(1 - p^k)^{n-k+1} \le \mathrm{P}(L_n(\mathrm{S}) < k) \le (1 - qp^k)^{n-k+1} \tag{3.3}$$

for $k = 1, \ldots, n$ and $0 < p < 1$. Muselli (2000) showed that, under the condition that $k \ge p/q$, the following inequalities hold:

$$(1 - p^k)^{((n-k)q/(1-p^k)^k)+1} \le \mathrm{P}(L_n(\mathrm{S}) < k) \le (1 - p^k)^{(n-k)q/(1-p^k)+1}. \tag{3.4}$$

The upper bound of (3.4) is slightly better than that of (3.3).

For Markov dependent trials, Godbole and Schaffner (1993) provided Stein–Chen-type inequalities for $P(L_n(S) < k)$. So far, to the best of the authors' knowledge, there are no satisfactory lower and upper bounds for $P(L_n < k)$.

4. Spectral analysis and large deviation approximation

Goncharov's approximation (1.5) for $P(L_n(S) < k)$ is intimately related to the upper bound in (3.3). This connection can be seen as follows.

For large n, taking $k = \lfloor \log_{1/p} n \rfloor + x,\ x = 0, \pm 1, \pm 2, \ldots$, the upper bound in (3.3) is

$$\begin{aligned}(1 - qp^k)^{n-k+1} &= (1 - qp^{\lfloor \log_{1/p} n \rfloor + x})^{n - \lfloor \log_{1/p} n \rfloor - x + 1} \\ &= (1 - \lambda_n / n)^{n - \lfloor \log_{1/p} n \rfloor - x + 1} \\ &\sim \exp\{-\lambda_n [1 + o(1)]\},\end{aligned}$$

where

$$\lambda_n = n(1-p)p^{\lfloor \log_{1/p} n \rfloor + x}.$$

This yields Goncharov's result. However, if k is fixed and n is very large, neither (1.5) nor (1.6) provides a good approximation for $P(L_n(S) < k)$ in terms of the relative error. To illustrate the relationship and connection, numerical examples are provided in the next section.

A square matrix $\boldsymbol{B} = [b_{ij}]$ is said to be nonnegative if $b_{ij} \geq 0$ for all i and j. If all $b_{ij} > 0$, we say that the matrix $\boldsymbol{B}$ is positive. A nonnegative matrix $\boldsymbol{B}$ is referred to as primitive if there exists a positive integer m such that $\boldsymbol{B}^m$ is positive. A nonnegative matrix $\boldsymbol{B}$ is referred to as a substochastic matrix if $0 < \sum_{j=1}^{l} b_{ij} \leq 1$ for all $i = 1, 2, \ldots, l$ and $0 < \sum_{j=1}^{l} b_{ij} < 1$ for at least one i.

If $\boldsymbol{N}$ is the square matrix defined by (2.3), let $\boldsymbol{N}^*$ denote the square matrix obtained by deleting the first row and the first column. In other words, the initial (dummy) state $\varnothing$ is deleted from the state space Ω and the imbedded Markov chain $\{Y_t\}$ has transition probability matrix $\boldsymbol{N}^*$ instead of $\boldsymbol{N}$. Let $\lambda_1, \ldots, \lambda_{2(k-1)}$ be the ordered eigenvalues of $\boldsymbol{N}^*$ in the sense that $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{2(k-1)}|$.

Theorem 4.1. (Perron–Frobenius.) *Given $k \geq 2$, if the transition probabilities p_{FF}, p_{FS}, p_{SF}, and p_{SS} are all greater than zero, then*

(i) *$\boldsymbol{N}^*$ is a primitive substochastic matrix and*

(ii) *the largest eigenvalue λ_1 satisfies $0 < \lambda_1 < 1$.*

Proof. It follows from the definition of $\boldsymbol{N}$ that the matrix $\boldsymbol{N}^*$ is nonnegative and substochastic. Taking $m = k + 1$, it is easy to see that, for $1 \leq i, j \leq k - 1$,

$$\begin{aligned}P(Y_{t+k+1} = S^j \mid Y_t = F^i) \geq P(&X_{t+1} = S, \ldots, X_{t+k-j} = S, X_{t+k-j+1} = F, \\ &X_{t+k-j+2} = S, \ldots, X_{t+k+1} = S \mid X_t = F) > 0\end{aligned}$$

for every t. Similarly, all other elements in $\boldsymbol{N}^*$ are positive. Hence, $\boldsymbol{N}^*$ is a primitive substochastic matrix. The result (ii) is exactly the so-called Perron–Frobenius eigenvalue theorem (see Seneta (1981)).

From the definition of $\boldsymbol{N}^*$, it is clear that the transition probability matrix $\boldsymbol{N}$ can always be decomposed into the following form:

$$\boldsymbol{N} = \left[\begin{array}{c|c} 0 & \boldsymbol{\xi}_0^* \\ \hline \boldsymbol{0} & \boldsymbol{N}^* \end{array}\right],$$

and $\boldsymbol{N}$ has the same $2(k-1)$ eigenvalues $\lambda_1, \ldots, \lambda_{2(k-1)}$ as $\boldsymbol{N}^*$, except that it has an additional eigenvalue $\lambda_{2k-1} = 0$ induced by the initial (dummy) state. Set $l = 2k-1$ and let $\boldsymbol{\eta}_1^\top, \ldots, \boldsymbol{\eta}_l^\top$ be eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_l$.

For simplicity of notation, in the following we present the results and proofs for the case that λ_1 has multiplicity one.

Note that the column vector $\mathbf{1}^\top$ can always be written as a linear combination of eigenvectors as

$$\mathbf{1}^\top = \sum_{i=1}^{l} c_i \boldsymbol{\eta}_i^\top. \tag{4.1}$$

It follows from Theorem 2.1 and (4.1) that

$$\mathrm{P}(L_n < k) = \sum_{i=1}^{l} c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top \lambda_i^n, \tag{4.2}$$

and the probability generating function for the sequence of probabilities $\{\mathrm{P}(L_n < k)\}_{n=1}^{\infty}$ is

$$\psi(t) = \sum_{i=1}^{l} \frac{c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top}{1 - \lambda_i t}.$$

In the following, we first provide a lemma required to prove our main result, and then derive the large deviation approximation for the tail probability $\mathrm{P}(L_n < k)$ for large n.

Lemma 4.1. *We have* $c_1^* := c_1 \boldsymbol{\xi}_0 \boldsymbol{\eta}_1^\top > 0$.

Proof. From the definition of L_n and (4.2), we have, for every n and $k \geq 2$,

$$\begin{aligned} 0 &< \mathrm{P}(X_1 = \mathrm{F}, X_2 = \mathrm{S}, X_3 = \mathrm{F}, \ldots, X_n = \mathrm{S}) \\ &\leq \mathrm{P}(L_n < k) = \lambda_1^n \left[c_1^* + \sum_{i=2}^{l} c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top \left(\frac{\lambda_i}{\lambda_1} \right)^n \right]. \end{aligned} \tag{4.3}$$

Since $|\lambda_i/\lambda_1|^n \to 0$ as $n \to \infty$ for $i = 2, 3, \ldots, l$, it follows that, as $n \to \infty$,

$$\sum_{i=2}^{l} |c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top| \left| \frac{\lambda_i}{\lambda_1} \right|^n \to 0. \tag{4.4}$$

Because c_1^* does not depend on n and $\lambda_1 > 0$, the result follows from (4.3) and (4.4).

Theorem 4.2. *Given k, the probability that L_n is less than k converges to zero exponentially in the following large deviation sense:*

$$\lim_{n \to \infty} \frac{1}{n} \log \mathrm{P}(L_n < k) = -\beta, \tag{4.5}$$

where $\beta = -\log \lambda_1$ is a positive constant.

In plain words, $\mathrm{P}(L_n < k)$ tends to zero with an exponential rate $\exp\{(n \log \lambda_1)(1 + o(1))\}$, i.e. for large n,

$$\mathrm{P}(L_n < k) \sim \exp(n \log \lambda_1).$$

Similarly, the above result also holds for the random variables $L_n(\mathrm{F})$ and $L_n(\mathrm{S})$:

$$\mathrm{P}(L_n(\mathrm{F}) < k) \sim \exp(n \log \lambda_1(0))$$

and

$$\mathrm{P}(L_n(\mathrm{S}) < k) \sim \exp(n \log \lambda_1(1)),$$

where $\lambda_1(0)$ and $\lambda_1(1)$ are the largest eigenvalues of the transition probability matrices $\boldsymbol{N}_0$ and $\boldsymbol{N}_1$, respectively.

Proof. For any arbitrarily small ε, since $(\lambda_i/\lambda_1)^n \to 0$ as $n \to \infty$ for $i = 2, 3, \ldots, l$, it follows that there exists an n_0 such that

$$\left| \sum_{i=2}^{l} \frac{c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top}{c_1^*} \left(\frac{\lambda_i}{\lambda_1} \right)^n \right| < \varepsilon$$

for all $n > n_0$. Since $c_1^* > 0$ as shown in Lemma 4.1, it follows from (4.3) that

$$\lambda_1^n (c_1^* - \varepsilon) \le \mathrm{P}(L_n < k) \le \lambda_1^n (c_1^* + \varepsilon)$$

for all $n > n_0$. The result (4.5) follows by applying log to the above inequalities, dividing by n, and letting $n \to \infty$.

Furthermore, the following stronger results also hold.

Theorem 4.3. (i) *We have*

$$\lim_{n \to \infty} \frac{\mathrm{P}(L_n < k)}{c_1^* \lambda_1^n} = 1.$$

(ii) *For large n,*

$$\mathrm{P}(L_n < k) = c_1^* \lambda_1^n (1 + O(\mathrm{e}^{-n\gamma})),$$

where $c_1^* = c_1 \boldsymbol{\xi}_0 \boldsymbol{\eta}_1^\top$ *and* $\gamma = \log |\lambda_1/\lambda_2|$.

Proof. From Lemma 4.1 and (4.3), we have

$$\mathrm{P}(L_n < k) = c_1^* \lambda_1^n \left[1 + \frac{1}{c_1^*} \sum_{i=2}^{l} c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top \left(\frac{\lambda_i}{\lambda_1} \right)^n \right].$$

By (4.4), there exists a $c_2^* > 0$ such that, for large n,

$$\left| \frac{1}{c_1^*} \sum_{i=2}^{l} c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top \left(\frac{\lambda_i}{\lambda_1} \right)^n \right| \le \frac{1}{c_1^*} \sum_{i=2}^{l} |c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top| \left| \frac{\lambda_2}{\lambda_1} \right|^n \le c_2^* \left| \frac{\lambda_2}{\lambda_1} \right|^n.$$

It follows that, for large n,

$$c_1^* \lambda_1^n \left[1 - c_2^* \left| \frac{\lambda_2}{\lambda_1} \right|^n \right] \le \mathrm{P}(L_n < k) \le c_1^* \lambda_1^n \left[1 + c_2^* \left| \frac{\lambda_2}{\lambda_1} \right|^n \right]. \tag{4.6}$$

Result (i) is a direct consequence of (4.6). In addition, since $c_2^* |\lambda_2/\lambda_1|^n = O(\mathrm{e}^{-n\gamma})$, where $\gamma = \log |\lambda_1/\lambda_2|$, result (ii) also follows from (4.6).

The results given in Theorems 4.2 and 4.3 also hold for $L_n(\mathrm{S})$ and $L_n(\mathrm{F})$. The proofs are the same. It is important to note that, for given k, the eigenvalues $\{\lambda_i\}_1^l$ and l are functions of k. Hence, our results depend on k. For example, in the case of $L_n(\mathrm{S})$ with $k = 1$, we have $l = 1$ and $\lambda_1 = q$, and therefore $\mathrm{P}(L_n(\mathrm{S}) < 1) = q^n$.

In view of Theorems 4.2 and 4.3, we expect that, for moderate or large n, the large deviation approximations will out perform other bounds and approximations mentioned in Section 3. In fact, the exponential rate $\beta = -\log \lambda_1$ and coefficient $c_1^*\lambda_1^n$ cannot be further improved. It is easy to see that for the general case where λ_1 has multiplicity v for some v with $1 < v \le l$, we only need to set $c_1^* = \sum_{i=1}^{v} c_i \boldsymbol{\xi}_0 \boldsymbol{\eta}_i^\top$ and all results remain unchanged and the proofs are the same.

5. Numerical results

In this section, our main goal is to present several numerical examples of exact distributions for the longest runs $L_n(\mathrm{F})$, $L_n(\mathrm{S})$, and L_n to illustrate our theoretical results. Numerical comparisons of exact distributions with various bounds and approximations are also given.

First, we compute the exact probabilities for $L_n(\mathrm{S})$ and L_n for various values of k. Table 1 presents the probabilities $\mathrm{P}(L_n(\mathrm{S}) < k)$ under the assumption that $\{X_t\}$ are i.i.d. two-state trials with $p = 0.2$ and 0.5 for $n = 20$, 50, 100, 1000, and 10 000. Table 2 gives the probabilities under the assumption that $\{X_t\}$ is a sequence of Markov dependent two-state trials with transition probability matrix

$$\boldsymbol{A}^* = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix} \tag{5.1}$$

and the initial probabilities $p = q = 0.5$. The distributions of $L_n(\mathrm{S})$ for $p = q = 0.5$ and $n = 20$, 30, 50, and 100 are shown graphically in Figure 1.

From Tables 1 and 2, it can be seen that, for given k, the exact probabilities $\mathrm{P}(L_n(\mathrm{S}) < k)$ are quite different in the i.i.d. and Markov dependent trials. For fixed k, all probabilities converge to zero exponentially. Figure 1 shows that the exact distributions of $L_n(\mathrm{S})$ for $n = 20$, 30, 50, 100

TABLE 1: Selected probabilities for i.i.d. two-state trials.

	$\mathrm{P}(L_n(\mathrm{S}) < 4)$		$\mathrm{P}(L_n(\mathrm{S}) < 8)$	
n	$p = 0.2$	$p = 0.5$	$p = 0.2$	$p = 0.5$
20	0.9781	0.5220	1.0000	0.9727
50	0.9410	0.1726	0.9999	0.9164
100	0.8823	0.0273	0.9998	0.8298
1000	0.2770	1.04×10^{-16}	0.9980	0.1389
10 000	2.57×10^{-6}	6.63×10^{-161}	0.9797	2.39×10^{-9}
	$\mathrm{P}(L_n < 4)$		$\mathrm{P}(L_n < 6)$	
n	$p = 0.2$	$p = 0.5$	$p = 0.2$	$p = 0.5$
20	0.0171	0.2316	0.1897	0.7631
50	1.83×10^{-5}	0.0188	0.0091	0.4559
100	2.06×10^{-10}	2.85×10^{-4}	5.76×10^{-5}	0.1932
1000	1.67×10^{-99}	5.15×10^{-37}	1.53×10^{-44}	3.75×10^{-8}
10 000	0	0	0	2.84×10^{-75}

TABLE 2: Selected probabilities for Markov dependent trials with transition probability matrix $\boldsymbol{A}^*$ defined in (5.1).

n	$\mathrm{P}(L_n(\mathrm{S}) < 4)$	$\mathrm{P}(L_n(\mathrm{S}) < 8)$	$\mathrm{P}(L_n < 4)$	$\mathrm{P}(L_n < 6)$
20	0.2094	0.7375	0.0063	0.1797
50	0.0165	0.4020	1.13×10^{-6}	0.0074
100	2.40×10^{-4}	0.1462	6.41×10^{-13}	3.57×10^{-5}
1000	2.01×10^{-37}	1.80×10^{-9}	2.30×10^{-125}	8.08×10^{-47}
10 000	0	1.55×10^{-88}	0	0

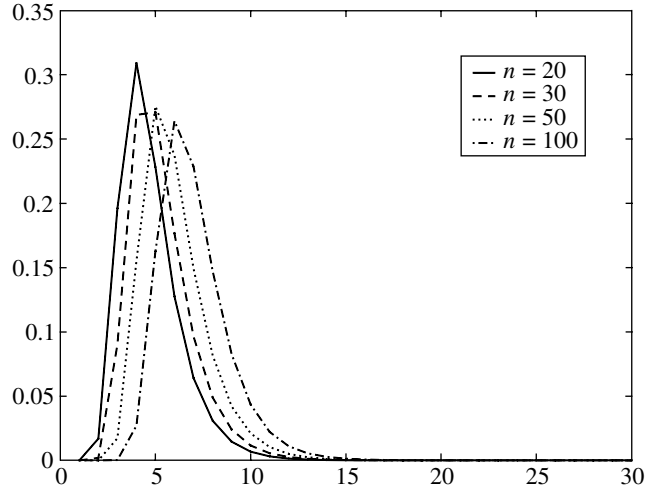

FIGURE 1: Distributions of $L_n(\mathrm{S})$ with $p = q = 0.5$.

are highly skewed to the right and $\mathrm{E}\, L_n(\mathrm{S})$ moves toward ∞ very slowly (approximately at rate $\log_{1/p} n$).

Next, we compare various bounds and approximations with exact probabilities for $L_n(\mathrm{S})$ and L_n. For simplicity, in Tables 3–6 we use the following notation:

(i) L_{S} and U_{S}: the Stein–Chen-type lower and upper bounds given by (3.2) with p and q interchanged;

(ii) L_{F} and U_{F}: the product-type lower and upper bounds given by (3.3);

(iii) L_{M} and U_{M}: Muselli's product-type lower and upper bounds given by (3.4);

(iv) A_{G}: Goncharov's approximation given by (1.5);

(v) A_{L}: large deviation approximation given by (4.5);

(vi) E: the exact probability given by (2.4) or (2.6).

Tables 3 and 4 compare the exact probabilities $\mathrm{P}(L_n(\mathrm{S}) < k)$ with their bounds and approximations under the i.i.d. assumption. Tables 5 and 6 compare the exact probabilities

TABLE 3: Comparisons of the exact $P(L_n(S) < k)$ with bounds and approximations for i.i.d. two-state trials with $p = 0.3$ and $k = 4, 8$.

n	L_S	L_F	L_M	A_L	E	A_G	U_F	U_M	U_S
					$k = 4$				
50	0.7288	0.6823	0.7568	0.7475	0.7590	0.7531	0.7655	0.7617	0.8033
100	0.5397	0.4543	0.5640	0.5588	0.5674	0.5672	0.5761	0.5717	0.6142
1000	−0.0338	0.0003	0.0028	0.0030	0.0030	0.0034	0.0035	0.0033	0.0408
10 000	−0.0373	4.89×10^{-36}	2.91×10^{-26}	5.27×10^{-26}	5.35×10^{-26}	2.37×10^{-25}	2.06×10^{-25}	1.20×10^{-25}	0.0373
					$k = 8$				
50	0.9974	0.9972	0.9980	0.9977	0.9980	0.9977	0.9980	0.9980	0.9987
100	0.9951	0.9939	0.9957	0.9954	0.9957	0.9954	0.9957	0.9957	0.9964
1000	0.9548	0.9369	0.9554	0.9551	0.9554	0.9551	0.9554	0.9554	0.9561
10 000	0.6313	0.5191	0.6318	0.6316	0.6318	0.6317	0.6319	0.6319	0.6326

TABLE 4: Comparisons of the exact $P(L_n(S) < k)$ with bounds and approximations for i.i.d. two-state trials with $p = 0.8$ and $k = 4, 6$.

n	L_S	L_F	L_M	A_L	E	A_G	U_F	U_M	U_S
					$k = 4$				
50	−0.2245	1.75×10^{-11}	2.77×10^{-18}	1.43×10^{-5}	2.28×10^{-5}	0.0166	0.0180	0.0002	0.2670
100	−0.2454	6.33×10^{-23}	4.05×10^{-37}	2.04×10^{-10}	3.26×10^{-10}	0.0003	0.0003	2.13×10^{-8}	0.2461
1000	−0.2458	6.81×10^{-229}	0.0000	1.23×10^{-97}	1.97×10^{-97}	2.65×10^{-36}	9.81×10^{-38}	3.60×10^{-78}	0.2458
					$k = 6$				
50	−0.2725	1.14×10^{-6}	4.66×10^{-8}	0.0064	0.0091	0.0727	0.0886	0.0196	0.4615
100	−0.3601	2.87×10^{-13}	3.07×10^{-16}	4.06×10^{-5}	5.83×10^{-5}	0.0053	0.0060	0.0003	0.3739
1000	−0.3670	4.28×10^{-132}	1.65×10^{-163}	1.21×10^{-44}	1.74×10^{-44}	1.70×10^{-23}	5.36×10^{-24}	1.98×10^{-36}	0.3670

TABLE 5: Comparisons of the exact $P(L_n(S) < k)$ with large deviation approximations for Markov dependent two-state trials with transition probability matrix $\boldsymbol{A}^*$.

	$k = 4$		$k = 8$	
n	E	A_L	E	A_L
20	0.2094	0.1841	0.7375	0.6673
50	0.0165	0.0145	0.4020	0.3637
100	2.40×10^{-4}	2.11×10^{-4}	0.1462	0.1323
1000	2.01×10^{-37}	1.77×10^{-37}	1.81×10^{-9}	1.64×10^{-9}
10 000	0	0	1.55×10^{-88}	1.40×10^{-88}

TABLE 6: Comparisons of the exact $P(L_n < k)$ with large deviation approximations for Markov dependent two-state trials with transition probability matrix $\boldsymbol{A}^*$.

	$k = 4$		$k = 8$	
n	E	A_L	E	A_L
20	0.9014	0.8868	0.9997	0.9995
50	0.7527	0.7405	0.9990	0.9989
100	0.5574	0.5484	0.9979	0.9977
1000	0.0025	0.0025	0.9775	0.9774
10 000	8.23×10^{-27}	8.09×10^{-27}	0.7955	0.7954

$P(L_n(S) < k)$ and $P(L_n < k)$ with their large deviation approximations under the assumption of Markov dependence.

In view of Table 3, for small p ($p = 0.3$) all the upper and lower bounds and approximations perform reasonably well except the Stein–Chen-type bounds. It can be shown that, for fixed k, the upper and lower Stein–Chen bounds tend to $\pm(2kq - 1)p^k$ as $n \to \infty$. For large p ($p = 0.8$), the performance of all the bounds and approximations is poor except for that of the probability large deviation approximations, which remains good. In view of Tables 3 and 4, the numerical results clearly suggest that U_F and A_G behave approximately the same with an exponential rate $-\log(1 - qp^k)$ while the upper bounds U_M are slightly better than U_F, having an exponential rate $-(q/(1 - p^k))\log(1 - p^k)$. It is not surprising that the large deviation approximations A_L perform extremely well in all cases in Tables 3–6 for moderate and large n, regardless of the value of p. For fixed k, the numerical results in Tables 3 and 4 clearly suggest that L_S, L_F, A_G, U_F, and U_S have the wrong exponential rates, tending to zero.

All numerical results are computed using a Pentium® III PC with 733 MHz CPU and standard hardware configuration. The computation times for all cases, including those for eigenvalues λ_1 in A_L, are within a few seconds. We have also computed exact probabilities E and the large deviation approximations A_L for large k and n. The case of $k = 100$ and $n = 10\,000$ took 0.33 second, whereas the case of $k = 500$ and $n = 10\,000$ took 33 seconds. Even for very large n ($\leq 500\,000$), computing the exact distribution (for all $k \sim \log n$) is easy and fast. Therefore, bounds and approximations are not recommended unless n and k are extremely large, such as $n \geq 10^{12}$. Given the rapid increase of computer speed nowadays, we believe that computation cost is not a problem in this type of computation. The MATLAB® programs for computing exact distributions, the large deviation approximations, and various bounds are available upon request from the first author of this paper.

6. Discussion and extension

Let us now consider a sequence $\{X_t\}$ of multistate Markov dependent trials with state space $\mathcal{F} = \{1, 2, \ldots, m\}$, $m \geq 3$. For $i \in \mathcal{F}$, let $L_n(i)$ be the length of the longest run of i and L_n be the length of the longest run of any state. All the results in Sections 2 and 4 can be extended to the multistate trials. The proofs of these results require only minor modifications. For example, given k, the probability of $L_n(i)$ being less than k can be cast in terms of the waiting time random variable $W(\Lambda(k, i)) > n$ in the following sense:

$$\mathrm{P}(L_n(i) < k) = \mathrm{P}(W(\Lambda(k, i)) > n) = \boldsymbol{\xi}_0 \boldsymbol{N}_i^n \mathbf{1}^\top,$$

where $\Lambda(k, i) = i^k$ and $\boldsymbol{N}_i$ is the essential part of $\boldsymbol{M}_i$, the transition probability matrix of the imbedded Markov chain $\{Y_t(i)\}$ corresponding to the waiting time $W(\Lambda(k, i))$. The imbedded Markov chain is defined on the state space

$$\Omega_i = \{\phi\} \cup \mathcal{F} \cup \mathcal{F}(\Lambda(k, i)) = \{\phi, 1, 2, \ldots, m, i^2, i^3, \ldots, i^{k-1}, \alpha_i\},$$

and the corresponding transition probability matrix is set up in the same way as in (2.2). Furthermore, the probability of L_n being less than k can be cast in terms of the waiting time random variable $W(\Lambda_k)$ being greater than n, where $\Lambda_k = \bigcup_{i=1}^m \Lambda(k, i)$ is a compound pattern generated by m simple patterns, $\Lambda(k, i) = i^k$, $i = 1, 2, \ldots, m$, of size k. Mathematically,

$$W(\Lambda_k) = \min_{1 \leq i \leq m} \{W(\Lambda(k, i))\}$$

and

$$\mathrm{P}(L_n < k) = \mathrm{P}(W(\Lambda_k) \geq n + 1).$$

The proofs of these are straightforward and we omit them.

For large n, the large deviation approximation formula, $\mathrm{P}(L_n < k) \sim \exp(-n\beta)$ with $\beta = -\log \lambda_1$, remains true. Note that λ_1 is the largest eigenvalue of $\boldsymbol{N}$, and $\boldsymbol{N}$ is the essential part of the transition probability matrix $\boldsymbol{M}$ of the imbedded Markov chain $\{Y_t\}$ associated with the waiting time of the compound pattern $\bigcup_{i=1}^m \Lambda(k, i)$. The proof is the same as that given in Section 4.

In conclusion, the finite Markov chain imbedding technique not only enables us to compute the exact distributions for $L_n(i)$ and L_n efficiently, but also provides excellent large deviation approximations.

Acknowledgements

This work was partially supported by the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the referee for many helpful comments and suggestions.

References

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

Barbour, A. D., Chryssaphinou, O. and Roos, M. (1995). Compound Poisson approximation in reliability theory. *IEEE Trans. Reliab.* **44,** 398–402.

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

Boutsikas, M. V. and Koutras, M. V. (2000). Reliability approximation for Markov chain imbeddable systems. *Methodology Comput. Appl. Prob.* **2,** 393–411.

BURR, E. J. AND CANE, G. (1961). Longest run of consecutive observations having a special attribute. *Biometrika* **48,** 461–465.

CHAO, M. T. AND FU, J. C. (1989). A limit theorem of certain repairable systems. *Ann. Inst. Statist. Math.* **41,** 809–818.

CHAO, M. T. AND FU, J. C. (1991). The reliability of a large series system under the Markov structure. *Adv. Appl. Prob.* **23,** 894–908.

CHRYSSAPHINOU, O. AND PAPASTAVRIDIS, S. G. (1990). Limit distribution for a consecutive-k-out-of-n: F system. *Adv. Appl. Prob.* **22,** 491–493.

DERMAN, C., LIBERMAN, G. J. AND ROSS, S. (1982). On the consecutive-k-out-of-n:F systems. *IEEE Trans. Reliab.* **31,** 57–63.

DOI, M. AND YAMAMOTO, E. (1998). On the joint distribution of runs in a sequence of multi-state trials. *Statist. Prob. Lett.* **39,** 133–141.

ERDŐS, P. AND RÉVÉSZ, P. (1975). On the length of the longest head run. In *Topics in Information Theory* (Colloq. Math. Soc. János Bolyai **16**), eds I. Csiszár and P. Elias, North-Holland, Amsterdam, pp. 219–228.

FU, J. C. (1986). Bounds for reliability of large consecutive-k-out-of-n: F systems with unequal component probabilities. *IEEE Trans. Reliab.* **35,** 316–319.

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

FU, J. C. AND CHANG, Y. M. (2002). On probability generating functions for waiting time distributions of compound patterns in a sequence of multistate trials. *J. Appl. Prob.* **39,** 70–80.

FU, J. C. AND KOUTRAS, M. V. (1994). Distribution theory of runs: a Markov chain approach. *J. Amer. Statist. Soc.* **89,** 1050–1058.

GIBBONS, J. D. (1971). *Nonparametric Statistical Inference*. McGraw-Hill, New York.

GODBOLE, A. P. AND SCHAFFNER, A. A. (1993). Improved Poisson approximations for word patterns. *Adv. Appl. Prob.* **25,** 334–347.

GONCHAROV, V. L. (1944). On the field of combinatory analysis. *Izv. Akad. Nauk. SSSR Ser. Mat.* **8,** 3–48 (in Russian). English translation: *Amer. Math. Soc. Transl.* **19** (1962), 1–46.

GORDON, L., SCHILLING, M. F. AND WATERMAN, M. S. (1986). An extreme value theory for long head runs. *Prob. Theory Relat. Fields* **72,** 279–287.

HIRANO, K. (1986). Some properties of the distributions of order k. In *Fibonacci Numbers and Their Applications* (Patras, 1984), Reidel, Dordrecht, pp. 43–53.

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. (1997). Sooner waiting time problems in a sequence of trinary trials. *J. Appl. Prob.* **34,** 593–609.

LOU, W. Y. W. (1996). On runs and longest run tests: a method of finite Markov chain imbedding. *J. Amer. Statist. Soc.* **91,** 1595–1601.

MUSELLI, M. (2000). New improved bounds for reliability of consecutive-k-out-of-n:F systems. *J. Appl. Prob.* **37,** 1164–1170.

PAPASTAVRIDIS, S. G. AND KOUTRAS, M. V. (1993). Bounds for reliability of consecutive-k-within-m-out-of-n systems. *IEEE Trans. Reliab.* **42,** 156–160.

PETROV, V. V. (1965). On the probabilities of large deviations for sums of independent random variables. *Theory Prob. Appl.* **10,** 287–298.

PHILIPPOU, A. N. AND MAKRI, F. S. (1986). Success runs and longest runs. *Statist. Prob. Lett.* **4,** 211–215.

RÉNYI, A. (1970). *Probability Theory*. Academic Kiadó, Budapest.

SENETA, E. (1981). *Nonnegative Matrices and Markov chains*, 2nd edn. Springer, New York.

SCHILLING, M. F. (1990). The longest run of heads. *College Math. J.* **21,** 196–207.

SUMAN, K. A. (1994). The longest run of any letter in a randomly generated word. In *Runs and Patterns in Probability: Selected Papers* (Math. Appl. **283**), eds A. P. Godbole and S. G. Papastavridis, Kluwer, Dordrecht, pp. 119–130.