Ann. Inst. Statist. Math.
Vol. 48, No. 3, 551–561 (1996)

DERIVATION OF THE PROBABILITY DISTRIBUTION FUNCTIONS FOR SUCCESSION QUOTA RANDOM VARIABLES

DEMETRIOS L. ANTZOULAKOS[1] AND ANDREAS N. PHILIPPOU[2]

[1] *Department of Economics, University of Crete, Perivolia-Rethimno 74100, Greece*
[2] *Department of Mathematics, University of Patras, Patras 26110, Greece*

(Received August 22, 1994; revised September 26, 1995)

Abstract. The probability distribution functions (pdf's) of the sooner and later waiting time random variables (rv's) for the succession quota problem (k successes and r failures) are derived presently in the case of a binary sequence of order k. The probability generating functions (pgf's) of the above rv's are then obtained directly from their pdf's. In the case of independent Bernoulli trials, expressions for the pdf's in terms of binomial coefficients are also established.

Key words and phrases: Succession quota, sooner and later problems, Bernoulli trials, binary sequence of order k, probability distribution function, probability generating function, longest run.

1. Introduction

Feller (1968) applied renewal theory to obtain the pgf of the distribution of the waiting time until a success run of length k or a failure run of length r occurs in a sequence of independent Bernoulli trials with common success probability p ($0 < p < 1$). Ebneshahrashoob and Sobel (1990) considered the same problem as a part of the Succession Quota (SQ) problem where one waits for a run of k successes or (and) a run of r failures whichever comes sooner (later). This situation is referred to as SQ sooner (later) waiting time problem. Generalizations of this problem have been considered by Ling (1990), Aki (1992), Ling and Low (1993), Aki and Hirano (1993), Balasubramanian *et al.* (1993) and Chryssaphinou *et al.* (1994). However, none of the above authors derived exact formulae for the pdf's of the SQ rv's.

In the present paper we derive exact formulae for the pdf's of the SQ sooner and later waiting time rv's in terms of multinomial coefficients, by means of a simple combinatorial argument, in the case of a binary sequence of order k (see Theorems 2.1 and 2.3) and hence in the case of independent Bernoulli trials. In the latter case, in particular, we also obtain formulae for the pdf's entirely in terms of binomial coefficients (see Theorems 3.1 and 3.2). The pgf's of the SQ sooner and later waiting time rv's are derived directly from their pdf's (see Theorems 2.2 and 2.4).

2. Pdf's and pgf's for the SQ sooner and later waiting time problems

Let W_S (W_L) be a rv denoting the waiting time for a run of k (≥ 2) successes or (and) a run of r (≥ 2) failures whichever comes sooner (later) in a binary sequence of order k. Aki (1992) obtained the pgf's of the rv's W_S and W_L, but no pdf's. In this section, we consider this problem, and derive exact formulae for the pdf's of W_S and W_L from which their pgf's follow. We recall first the definition of a binary sequence of order k from Aki (1985).

DEFINITION 2.1. An infinite sequence $\{X_n\}_{n=0}^{\infty}$ of $\{0,1\}$-valued rv's is said to be a binary sequence of order k if there exists a positive integer k and k real numbers $0 < p_1, p_2, \ldots, p_k < 1$ such that

(1) $X_0 = 0$ almost surely, and

(2) $P(X_n = 1 \mid X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}) = p_j$, $n \geq 1$,

where $j = m - [(m-1)/k] \cdot k$, m is the smallest positive integer which satisfies $x_{n-m} = 0$, and $[x]$ denotes the greatest integer in x.

The outcomes "1" and "0" are usually called "success" and "failure", respectively.

THEOREM 2.1. *Let W_S be a rv denoting the waiting time until the occurrence of a success run of length k or a failure run of length r in the binary sequence $\{X_n\}_{n=1}^{\infty}$ of order k. Then for $w \geq \min\{k, r\}$*

$$\begin{aligned}
P(W_S = w) &= q_1^{r-1} \sum_{n=0}^{k-1} (p_0 p_1 \cdots p_n q_{n+1}) \sum_{1} \frac{(\sum_i \sum_j x_{ij})!}{\prod_i \prod_j x_{ij}!} \\
&\quad \times \prod_i \prod_j (q_1^{i-1} p_1 p_2 \cdots p_j q_{j+1})^{x_{ij}} \\
&\quad + (p_1 p_2 \cdots p_k) \sum_{n=0}^{r-1} q_1^n \sum_{2} \frac{(\sum_i \sum_j y_{ij})!}{\prod_i \prod_j y_{ij}!} \\
&\quad \times \prod_i \prod_j (p_1 p_2 \cdots p_i q_{i+1} q_1^{j-1})^{y_{ij}},
\end{aligned}$$

where $q_l = 1 - p_l$ ($1 \leq l \leq k$), $p_0 = 1$, and the inner summations $\sum_1$ and $\sum_2$ are taken over all non-negative integers x_{ij} and y_{ij} satisfying the conditions $n + \sum_i \sum_j (i+j) x_{ij} = w - r$ for $1 \leq i \leq r-1$ and $1 \leq j \leq k-1$, and $n + \sum_i \sum_j (i+j) y_{ij} = w - k$ for $1 \leq i \leq k-1$ and $1 \leq j \leq r-1$, respectively.

PROOF. Let $W_S^{(0)}$ (resp. $W_S^{(1)}$) be a rv denoting the waiting time until a failure (resp. success) run of length r (resp. k) occurs before a success (resp. failure) run of length k (resp. r) in the binary sequence $\{X_n\}_{n=1}^{\infty}$ of order k. Then

$$P(W_S = w) = P(W_S^{(0)} = w) + P(W_S^{(1)} = w), \qquad w \geq \min\{k, r\}. \tag{2.1}$$

We shall first derive $P(W_S^{(0)} = w)$, $w \geq r$. Modifying a combinatorial argument of Philippou and Muwafi (1982), we note that a typical element of the event $(W_S^{(0)} = w)$ is an arrangement

$$\underbrace{11\ldots1}_{n}\alpha_1\alpha_2\ldots\alpha_{\sum_i\sum_j x_{ij}}\underbrace{00\ldots0}_{r}, \qquad 0 \leq n \leq k-1, \tag{2.2}$$

of the numbers "0" and "1", such that x_{ij} of the α's are of the form $\underbrace{00\ldots0}_{i}\underbrace{11\ldots1}_{j}$, $1 \leq i \leq r-1$ and $1 \leq j \leq k-1$, and $n + \sum_i\sum_j(i+j)x_{ij} + r = w$. Fix n and x_{ij} ($1 \leq i \leq r-1$ and $1 \leq j \leq k-1$). Then the number of the above arrangements is

$$\frac{(\sum_i\sum_j x_{ij})!}{\prod_i\prod_j x_{ij}!}. \tag{2.3}$$

Observe that (i) the probability of the pattern $\underbrace{11\ldots1}_{n}$ is $p_0p_1\cdots p_n$ ($0 \leq n \leq k-1$) and determines probability equal to q_{n+1} for the next first occurring "0"; (ii) the probability of the pattern $\underbrace{00\ldots0}_{i}\underbrace{11\ldots1}_{j}$ is $q_1^{i-1}p_1p_2\cdots p_j$ multiplied by the probability of its first "0", and determines probability equal to q_{j+1} for the next first occurring "0"; (iii) the probability of the pattern $\underbrace{00\ldots0}_{r}$ is q_1^{r-1} multiplied by the probability of its first "0". Since there are $1 + \sum_i\sum_j x_{ij}$ first "0"s, one with probability q_{n+1} and $\sum_i x_{ij}$ with probability q_{j+1}, $1 \leq j \leq k-1$, the probability of each one of the above arrangements is

$$(p_0p_1\cdots p_nq_{n+1})\prod_i\prod_j(q_1^{i-1}p_1p_2\cdots p_jq_{j+1})^{x_{ij}}q_1^{r-1}. \tag{2.4}$$

From (2.2)–(2.4) it follows that

$$\begin{aligned} P(W_S^{(0)} = w) = q_1^{r-1}\sum_{n=0}^{k-1}(p_0p_1\cdots p_nq_{n+1})\sum_1\frac{(\sum_i\sum_j x_{ij})!}{\prod_i\prod_j x_{ij}!} \\ \times\prod_i\prod_j(q_1^{i-1}p_1p_2\cdots p_jq_{j+1})^{x_{ij}}, \qquad w \geq r. \end{aligned} \tag{2.5}$$

Replacing in (2.2) "0", "1", r, k and x_{ij} by "1", "0", k, r and y_{ij}, respectively, and following the same procedure we get

$$\begin{aligned} P(W_S^{(1)} = w) = (p_1p_2\cdots p_k)\sum_{n=0}^{r-1}q_1^n\sum_2\frac{(\sum_i\sum_j y_{ij})!}{\prod_i\prod_j y_{ij}!} \\ \times\prod_i\prod_j(p_1p_2\cdots p_iq_{i+1}q_1^{j-1})^{y_{ij}}, \qquad w \geq k. \end{aligned} \tag{2.6}$$

The proof of the theorem follows by means of (2.1), (2.5) and (2.6).

For $k = r = 2$ and $p_1 = p_2 = p$, Theorem 2.1 reduces to Theorem 4.1 of Ling (1990).

Next, we obtain the pgf of the rv W_S directly from its pdf.

THEOREM 2.2. *Let W_S be as in Theorem 2.1 and denote by $G_S(t)$ its pgf. Then*

$$G_S(t) = \left(A_{r-1}(t) \sum_{j=0}^{k-1} B_j(t) + B(t) \sum_{i=0}^{r-1} A_i(t) \right) \times \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1},$$

where $A_m(t) = (q_1 t)^m$, $B_m(t) = q_{m+1} t^{m+1} \prod_{i=0}^{m} p_i$ and $B(t) = t^k \prod_{i=1}^{k} p_i$.

PROOF. Let $W_S^{(0)}$ and $W_S^{(1)}$ be as in the proof of Theorem 2.1, and let $G_S^{(0)}(t)$ and $G_S^{(1)}(t)$ be their pgf's, respectively. Then

$$G_S(t) = G_S^{(0)}(t) + G_S^{(1)}(t). \tag{2.7}$$

We shall first derive $G_S^{(0)}(t)$. From (2.5) it follows that

$$G_S^{(0)}(t) = A_{r-1}(t) \sum_{n=0}^{k-1} B_n(t) \sum_{w=0}^{\infty} \sum_{*} \frac{(\sum_i \sum_j x_{ij})!}{\prod_i \prod_j x_{ij}!} \prod_i \prod_j (A_{i-1}(t) B_j(t))^{x_{ij}},$$

where the inner summation $\sum_*$ is taken over all non-negative integers x_{ij} satisfying the condition $n + \sum_i \sum_j (i+j) x_{ij} = w$, $1 \leq i \leq r-1$ and $1 \leq j \leq k-1$. Let C_n $(0 \leq n \leq k-1)$ be the set of non-negative solutions of the above equation for $0 \leq w < \infty$. We observe that $C_n = C$, where C is the set of non-negative solutions of the equation $\sum_i \sum_j x_{ij} = w$, $1 \leq i \leq r-1$ and $1 \leq j \leq k-1$, for $0 \leq w < \infty$. Thus, by the multinomial theorem we get

$$G_S^{(0)}(t) = A_{r-1}(t) \sum_{j=0}^{k-1} B_j(t) \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1}. \tag{2.8}$$

Using the same procedure and (2.6) we get

$$G_S^{(1)}(t) = B(t) \sum_{i=0}^{r-1} A_i(t) \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1}. \tag{2.9}$$

The proof of the theorem follows by means of (2.7)–(2.9).

We proceed now to the SQ later waiting time problem.

THEOREM 2.3. *Let W_L be a rv denoting the waiting time until the occurrence of a success run of length k and a failure run of length r in the binary sequence $\{X_n\}_{n=1}^{\infty}$ of order k. Then for $w \geq k+r$*

$$P(W_L = w) = q_1^{r-1} \sum_{n=0}^{k-1} (p_0 p_1 \cdots p_n q_{n+1}) \sum_3 \frac{(x_k + \sum_i \sum_j x_{ij})!}{x_k! \prod_i \prod_j x_{ij}!} \times \prod_i \prod_j (q_1^{i-1} p_1 p_2 \cdots p_j q_{j+1})^{x_{ij}} (p_1 p_2 \cdots p_k)^{x_k} + (p_1 p_2 \cdots p_k) \sum_{n=0}^{r-1} q_1^n \sum_4 \frac{(y_r + \sum_i \sum_j y_{ij})!}{y_r! \prod_i \prod_j y_{ij}!} \times \prod_i \prod_j (p_1 p_2 \cdots p_i q_{i+1} q_1^{j-1})^{y_{ij}} (q_1^r)^{y_r},$$

where $q_l = 1 - p_l$ $(1 \leq l \leq k)$, $p_0 = 1$, $q_{k+1} = q_1$, the summation $\sum_3$ is taken over all non-negative integers x_{ij} and x_k satisfying the condition $n + \sum_i \sum_j (i+j) x_{ij} + k x_k = w - r$ $(1 \leq i \leq r-1$ and $1 \leq j \leq k)$ and for at least one i either $x_{ik} \geq 1$, or $x_k \geq 1$, and the summation $\sum_4$ is taken over all non-negative integers y_{ij} and y_r satisfying the condition $n + \sum_i \sum_j (i+j) y_{ij} + r y_r = w - k$ $(1 \leq i \leq k-1$ and $1 \leq j \leq r)$ and for at least one i either $y_{ir} \geq 1$, or $y_r \geq 1$.

PROOF. Let $W_L^{(0)}$ (resp. $W_L^{(1)}$) be a rv denoting the waiting time until a failure (resp. success) run of length r (resp. k) occurs later than a success (resp. failure) run of length k (resp. r) in the binary sequence $\{X_n\}_{n=1}^{\infty}$ of order k. Then

$$P(W_L = w) = P(W_L^{(0)} = w) + P(W_L^{(1)} = w), \qquad w \geq k+r. \tag{2.10}$$

We shall first derive $P(W_L^{(0)} = w)$, $w \geq k+r$. We observe that a typical element of the event $(W_L^{(0)} = w)$ is an arrangement

$$\underbrace{11\ldots1}_{n} \alpha_1 \alpha_2 \ldots \alpha_{\sum_i \sum_j x_{ij} + x_k} \underbrace{00\ldots0}_{r}, \qquad 0 \leq n \leq k-1,$$

of the numbers "0" and "1", such that x_{ij} of the α's are of the form $\underbrace{00\ldots0}_{i}\underbrace{11\ldots1}_{j}$, $1 \leq i \leq r-1$ and $1 \leq j \leq k$, x_k of the α's are of the form $\underbrace{11\ldots1}_{k}$, $n + \sum_i \sum_j (i+j) x_{ij} + k x_k = w - r$, and for at least one i either $x_{ik} \geq 1$, or $x_k \geq 1$. By assigning probabilities to the appearing patterns as in Theorem 2.1, and by the fact that the probability of any occurring pattern of the form $\underbrace{11\ldots1}_{k}$ is always $p_1 p_2 \cdots p_k$, we get

$$P(W_L^{(0)} = w) = q_1^{r-1} \sum_{n=0}^{k-1} (p_0 p_1 \cdots p_n q_{n+1}) \tag{2.11}$$

$$\times \sum_{3} \frac{(x_k + \sum_i \sum_j x_{ij})!}{x_k! \prod_i \prod_j x_{ij}!}$$
$$\times \prod_i \prod_j (q_1^{i-1} p_1 p_2 \cdots p_j q_{j+1})^{x_{ij}} (p_1 p_2 \cdots p_k)^{x_k},$$
$$w \geq k + r.$$

A similar argument leads, for $w \geq k + r$, to

$$P(W_L^{(1)} = w) = (p_1 p_2 \cdots p_k) \sum_{n=0}^{r-1} q_1^n \sum_{4} \frac{(y_r + \sum_i \sum_j y_{ij})!}{y_r! \prod_i \prod_j y_{ij}!} \tag{2.12}$$
$$\times \prod_i \prod_j (p_1 p_2 \cdots p_i q_{i+1} q_1^{j-1})^{y_{ij}} (q_1^r)^{y_r}.$$

The proof of the theorem follows by means of (2.10)–(2.12).

Next, we obtain the pgf of the rv W_L directly from its pdf. Our formula appears to be simpler than the corresponding one in Aki (1992).

THEOREM 2.4. *Let $A_m(t)$, $B_m(t)$ and $B(t)$ be as in Theorem 2.2, and denote by $G_L(t)$ the pgf of the rv W_L. Then*

$$G_L(t) = A_{r-1}(t) \sum_{j=0}^{k-1} B_j(t) \left\{ \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k} B_j(t) \right) - B(t) \right]^{-1} - \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1} \right\}$$
$$+ B(t) \sum_{i=0}^{r-1} A_i(t) \left\{ \left[1 - \left(\sum_{i=0}^{r-1} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) - A_r(t) \right]^{-1} - \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1} \right\}.$$

PROOF. Let $W_L^{(0)}$ and $W_L^{(1)}$ be as in the proof of Theorem 2.3, and let $G_L^{(0)}(t)$ and $G_L^{(1)}(t)$ be their pgf's, respectively. Then

$$G_L(t) = G_L^{(0)}(t) + G_L^{(1)}(t).$$

We shall first derive $G_L^{(0)}(t)$. From (2.11) it follows that

$$G_L^{(0)}(t) = A_{r-1}(t) \sum_{n=0}^{k-1} B_n(t) \sum_{w=0}^{\infty} \sum_{*} \frac{(x_k + \sum_i \sum_j x_{ij})!}{x_k! \prod_i \prod_j x_{ij}!}$$
$$\times \prod_i \prod_j (A_{i-1}(t) B_j(t))^{x_{ij}} (B(t))^{x_k},$$

where the inner summation $\sum_*$ is taken over all non-negative integers x_{ij} and x_k satisfying the condition $n + \sum_i \sum_j (i+j)x_{ij} + kx_k = w + k$, $1 \le i \le r-1$ and $1 \le j \le k$, and for at least one i either $x_{ik} \ge 1$, or $x_k \ge 1$. Let D_n $(0 \le n \le k-1)$ be the set of non-negative solutions of the above equation with the restrictions satisfied, for $0 \le w < \infty$. We observe that $D_n = E_1 - E_2$, where E_1 is the set of non-negative solutions of the equation $\sum_i \sum_j x_{ij} + x_k = w$, for $0 \le w < \infty$, and E_2 is the set of non-negative solutions of the same equation having $x_{ik} = 0$ $(1 \le i \le r-1)$ and $x_k = 0$, for $0 \le w < \infty$. Thus, by the multinomial theorem we get

$$G_L^{(0)}(t) = A_{r-1}(t) \sum_{j=0}^{k-1} B_j(t) \left\{ \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k} B_j(t) \right) - B(t) \right]^{-1} - \left[1 - \left(\sum_{i=0}^{r-2} A_i(t) \right) \left(\sum_{j=1}^{k-1} B_j(t) \right) \right]^{-1} \right\}.$$

A similar argument leads to $G_L^{(1)}(t)$ which completes the proof of the theorem.

Remark 1. For $p_1 = p_2 = \cdots = p_k = p$, Theorems 2.1 and 2.3 give the pdf's of W_S and W_L, respectively, in the case of independent Bernoulli trials with success probability p, and they are new results. We also note that $W_S^{(0)}$, $W_S^{(1)}$, $W_L^{(0)}$ and $W_L^{(1)}$ may not be usual real rv's in the sense that they may take the value ∞. Therefore, their pdf's and pgf's may not be proper.

3. Pdf's in terms of binomial coefficients for the SQ sooner and later waiting time problems: Bernoulli trials

In the present section, we derive expressions in terms of binomial coefficients for the pdf's of the SQ sooner and later waiting time rv's in the case of independent Bernoulli trials. First, we state two well-known results on occupancy problems (see, e.g. p. 105 of Riordan (1958)).

LEMMA 3.1. *Let $M(n,m,r)$ be the number of ways of distributing n identical balls into m different cells with no cell containing more than r balls. Then*

$$M(n,m,r) = \sum_{j=0}^{[n/(r+1)]} (-1)^j \binom{m}{j} \binom{n - j(r+1) + m - 1}{m-1}.$$

LEMMA 3.2. *Let $Q(n,m,r)$ be the number of ways of distributing n identical balls into m different cells with no cell containing more than r balls and no cell is empty. Then*

$$Q(n,m,r) = \sum_{j=0}^{[(n-m)/r]} (-1)^j \binom{m}{j} \binom{n - jr - 1}{m-1}.$$

Next, we obtain a new result which is very important for our derivations.

LEMMA 3.3. *Let* $\{i_1, i_2, \dots, i_m\}$ *be an* m*-combination of* $\{1, 2, \dots, n\}$. *Assume that* $1 \le i_1 < i_2 < \cdots < i_m \le n$ *and let* $C(n, m, r)$ *be the number of these combinations where* $i_1 \ge r+1$ *or* $i_s - i_{s-1} \ge r$ *for at least one* s $(2 \le s \le m)$ *or* $n - i_m \ge r - 1$. *Then*

$$C(n,m,r) = \sum_{j=0}^{m} (-1)^j \left\{ \binom{m}{j} \binom{n-r-j(r-1)}{m} + \binom{m}{j+1} \binom{n-r+1-j(r-1)}{m} \right\}.$$

PROOF. Let $j_1 = i_1$, $j_2 = i_2 - i_1$, $j_3 = i_3 - i_2, \dots, j_m = i_m - i_{m-1}$ and $j_{m+1} = n - i_m$. Then $C(n, m, r)$ coincides with the number of integral solutions of the equation

$$j_1 + j_2 + \cdots + j_{m+1} = n, \tag{3.1}$$

where $j_1 \ge r+1$ or $j_s \ge r$ for at least one s $(2 \le s \le m)$ or $j_{m+1} \ge r-1$, and it holds true that in any case $j_1, j_2, \dots, j_m \ge 1$ and $j_{m+1} \ge 0$. Let E_l $(1 \le l \le m+1)$ be the set of integral solutions of (3.1) satisfying the restriction for the respective j_l, and let $N(E_l)$ be the number of these solutions.

Consider intersections of the form $E_{t_1} \cap E_{t_2} \cap \cdots \cap E_{t_h}$. Taking cases according to whether E_1 or E_{m+1} belong to an intersection, and noting that the number of integral solutions of the equation $x_1 + x_2 + \cdots + x_k = n$, where $x_i \ge d_i$ $(1 \le i \le k)$ and $d = d_1 + d_2 + \cdots + d_k$, is $\binom{k+n-d-1}{k-1}$ (see, e.g. Berman and Fryer (1972), p. 76), we get

$$\begin{aligned} &N(\text{all intersections of the form } E_{t_1} \cap E_{t_2} \cap \cdots \cap E_{t_h}) \\ &\quad = \binom{m-1}{h-1}\binom{n-1-h(r-1)}{m} + \binom{m-1}{h-1}\binom{n-h(r-1)}{m} \\ &\qquad + \binom{m-1}{h-2}\binom{n-1-h(r-1)}{m} + \binom{m-1}{h}\binom{n-h(r-1)}{m} \\ &\quad = \binom{m}{h-1}\binom{n-1-h(r-1)}{m} + \binom{m}{h}\binom{n-h(r-1)}{m}. \end{aligned}$$

The proof of the lemma then follows by noting that $C(n, m, r) = N(E_1 \cup E_2 \cup \cdots \cup E_{m+1})$ and using the inclusion-exclusion principle.

We are now ready to express the pdf's of the SQ sooner and later waiting time rv's in terms of binomial coefficients.

THEOREM 3.1. *Let* $\tilde{W}_S$ *be a rv denoting the waiting time until the occurrence of a success run of length* k *or a failure run of length* r *in a sequence of independent*

Bernoulli trials with common success probability p $(0 < p < 1)$, *and set* $q = 1 - p$. *Also let* $M(n,m,r)$, $Q(n,m,r)$ *and* $C(n,m,r)$ *be as in Lemmas* 3.1, 3.2 *and* 3.3, *respectively, and set*

$$\begin{aligned} P_S(w;q,r,k) &= \zeta(r,w)\zeta(w,r+k-1)q^r p^{w-r} \\ &\quad + \zeta(r+2,w)\sum_{n=1}^{w-r-1} p^n q^{w-n} \\ &\quad \times \Bigg[M(w-r-n,n,r-1) \\ &\qquad - \zeta(k,n)\sum_{m=1}^{n-k+1} C(n,m,k)Q(w-r-n,m,r-1)\Bigg], \end{aligned}$$

where $\zeta(\cdot,\cdot)$ *is the zeta function defined by* $\zeta(u,v) = 1$ *if* $v \geq u$ *and* 0 *otherwise. Then for* $w \geq \min\{k,r\}$

$$P(\tilde{W}_S = w) = P_S(w;q,r,k) + P_S(w;p,k,r).$$

PROOF. Let $\tilde{W}_S^{(0)}$ (resp. $\tilde{W}_S^{(1)}$) be a rv denoting the waiting time until a failure (resp. success) run of length r (resp. k) occurs before a success (resp. failure) run of length k (resp. r). We shall first derive $P(\tilde{W}_S^{(0)} = w)$, $w \geq r$. Let $L_w^{(0)}$, $L_w^{(1)}$, S_w and X_w denote, respectively, the length of the longest failure run, the length of the longest success run, the number of successes and the outcome of the w-th trial in w independent Bernoulli trials with common success probability p. It is obvious that

$$P(\tilde{W}_S^{(0)} = r) = q^r = P_S(r;q,r,k). \tag{3.2}$$

For $w \geq r+1$, we observe that

$$\begin{aligned} &P(\tilde{W}_S^{(0)} = w) \\ &= q^r \sum_{n=1}^{w-r} P(L_{w-r}^{(0)} \leq r-1, L_{w-r}^{(1)} \leq k-1, S_{w-r} = n, X_{w-r} = 1) \\ &= q^r \Bigg\{\zeta(r,w)\zeta(w,r+k-1)p^{w-r} + \zeta(r+2,w)\sum_{n=1}^{w-r-1} q^{w-r-n}p^n \\ &\qquad \times N(L_{w-r}^{(0)} \leq r-1, L_{w-r}^{(1)} \leq k-1, S_{w-r} = n, X_{w-r} = 1)\Bigg\} \\ &= \zeta(r,w)\zeta(w,r+k-1)q^r p^{w-r} + \zeta(r+2,w)\sum_{n=1}^{w-r-1} q^{w-n}p^n \\ &\quad \times [N(L_{w-r}^{(0)} \leq r-1, S_{w-r} = n, X_{w-r} = 1) \\ &\qquad - \zeta(k,n)N(L_{w-r}^{(0)} \leq r-1, L_{w-r}^{(1)} \geq k, S_{w-r} = n, X_{w-r} = 1)]. \end{aligned} \tag{3.3}$$

It is not difficult to realize, by Lemmas 3.1, 3.2 and 3.3, respectively, that

$$(3.4)\qquad N(L_{w-r}^{(0)} \le r-1, S_{w-r}=n, X_{w-r}=1) = M(w-r-n, n, r-1),$$

and

$$(3.5)\qquad N(L_{w-r}^{(0)} \le r-1, L_{w-r}^{(1)} \ge k, S_{w-r}=n, X_{w-r}=1) = \sum_{m=1}^{n-k+1} C(n,m,k)Q(w-r-n,m,r-1).$$

From (3.2)–(3.5) it follows that $P(\tilde{W}_S^{(0)}=w) = P_S(w;q,r,k)$, $w \ge r$. By symmetry, we have that $P(\tilde{W}_S^{(1)}=w) = P_S(w;p,k,r)$, $w \ge k$, and this completes the proof of the theorem.

Remark 2. The methodology employed by Bradley ((1968), p. 258), may be modified in order to derive an alternative formula for $P(\tilde{W}_S = w)$.

THEOREM 3.2. *Let $\tilde{W}_L$ be a rv denoting the waiting time until the occurrence of a success run of length k and a failure run of length r in a sequence of independent Bernoulli trials with common success probability p $(0<p<1)$, and set $q=1-p$. Also let $Q(n,m,r)$ and $C(n,m,r)$ be as in Lemmas* 3.2 *and* 3.3, *respectively, and set*

$$P_L(w;q,r,k) = q^r p^{w-r} + \zeta(r+k+1,w)\sum_{n=k}^{w-r-1} p^n q^{w-n} \times \sum_{m=1}^{n-k+1} C(n,m,k)Q(w-r-n,m,r-1).$$

Then for $w \ge k+r$

$$P(\tilde{W}_L = w) = P_L(w;q,r,k) + P_L(w;p,k,r).$$

PROOF. Let $\tilde{W}_L^{(0)}$ (resp. $\tilde{W}_L^{(1)}$) be a rv denoting the waiting time until a failure (resp. success) run of length r (resp. k) occurs later than a success (resp. failure) run of length k (resp. r). Noting that

$$P(\tilde{W}_L^{(0)}=w) = q^r \sum_{n=k}^{w-r} P(L_{w-r}^{(0)} \le r-1, L_{w-r}^{(1)} \ge k, S_{w-r}=n, X_{w-r}=1), \qquad w \ge k+r,$$

and using (3.5) we have that $P(\tilde{W}_L^{(0)}=w) = P_L(w;q,r,k)$, $w \ge k+r$. By symmetry, we have that $P(\tilde{W}_L^{(1)}=w) = P_L(w;p,k,r)$, $w \ge k+r$, and this completes the proof of the theorem.

Acknowledgements

The authors wish to thank the referees for their helpful comments which led to improvements of the paper.

REFERENCES

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

Aki, S. (1992). Waiting time problems for a sequence of discrete random variables, *Ann. Inst. Statist. Math.*, **44**, 363–378.

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

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

Berman, G. and Fryer, K. D. (1972). *Introduction to Combinatorics*, Academic Press, New York.

Bradley, J. V. (1968). *Distribution-Free Statistical Tests*, Prentice-Hall, Englewood, Cliffs.

Chryssaphinou, O., Papastavridis, S. and Tsapelas, T. (1994). On the waiting time of appearance of given patterns, *Runs and Patterns in Probability* (eds. A. P. Godbole and S. G. Papastavridis), 231–241, Kluwer, Dordrecht.

Ebneshahrashoob, M. and Sobel, M. (1990). Sooner and later waiting time problems for Bernoulli trials: frequency and run quotas, *Statist. Probab. Lett.*, **9**, 5–11.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., Wiley, New York.

Ling, K. D. (1990). On geometric distributions of order $(k_1, \dots, k_m)$, *Statist. Probab. Lett.*, **9**, 163–171.

Ling, K. D. and Low, T. Y. (1993). On the soonest and latest waiting time distributions: succession quotas, *Comm. Statist. Theory Methods*, **22**, 2207–2221.

Philippou, A. N. and Muwafi, A. A. (1982). Waiting for the k-th consecutive success and the Fibonacci sequence of order k, *Fibonacci Quart.*, **20**, 28–32.

Riordan, J. (1958). *An Introduction to Combinatorial Analysis*, Wiley, New York.