A direct method to obtain the joint distribution of successes, failures and patterns in enumeration problems

Demetrios L. Antzoulakos and Michael V. Boutsikas

Department of Statistics and Insurance Science,

University of Piraeus, Piraeus 18534, Greece

Abstract. Let $\mathcal{E}$ be a pattern (single or composite) defined over a sequence of Bernoulli trials. Let X_n, S_n and F_n be the number of occurrences of $\mathcal{E}$, the number of successes and the number of failures in n trials, respectively. Also let T_r, S_{T_r} and F_{T_r} be the waiting time for the rth occurrence of $\mathcal{E}$ in an infinite sequence of trials, the number of successes and the number of failures observed at that time, respectively. In this article we present some simple results revealing that, if we know either the distribution of X_n or the distribution of T_r we can instantly derive the joint distribution of (X_n, S_n, F_n) and (T_r, S_{T_r}, F_{T_r}).

Key words and phrases: Patterns, success runs, scans, probability generating function, double generating function, waiting time distribution.

1. Introduction

In several areas of applied science it is quite common to get interested in the number X_n of occurrences of a single or composite pattern $\mathcal{E}$ in a sequence of $n \geq 1$ binary trials as well as in the waiting time T_r for the rth ($r \geq 1$) occurrence of $\mathcal{E}$ in an infinite sequence of outputs. Success runs and scans (strings of trials with a prespecified number of successes among them), are the two most widely used patterns for performing the necessary statistical analysis in many scientific areas, which include quality control, reliability theory, educational psychology, molecular biology, meteorology, actuarial science etc. During the last decades an upsurge of interest has been observed in the study of runs and scans, as the extensive recent literature indicates (see e.g. Glaz and Balakrishnan (1999), Glaz, Naus and Wallenstein (2001), Balakrishnan and Koutras (2002), and Fu and Lou (2003)).

When studying the number X_n of occurrences of a pattern $\mathcal{E}$ in a sequence of n binary trials, the number S_n of successes and/or the number F_n of failures appeared in the observed sequence provide useful information for the efficient study of the statistical model. Similarly, when studying the waiting time T_r for the rth occurrence of $\mathcal{E}$ in an infinite sequence of trials, information about the number S_{T_r} of successes and/or the number F_{T_r} of failures observed at that time could be of great practical importance. Aki and Hirano (1994) seem to have been the first to become attracted to these kinds of problems. More specifically, they proceeded to the investigation of the distribution of S_{T_1} and the joint distribution of (T_1, F_{T_1}) in the case where T_1 denotes the waiting time for the first occurrence of a success run of length k. Several extensions and variations of their waiting time model were subsequently studied by Aki and Hirano (1995), Antzoulakos and Philippou (1996), Balakrishnan (1997), Uchida (1998) and Chadjiconstantinidis and Koutras (2001). Recently, Chadjiconstantinidis *et al.* (2000) presented a unified approach for the study of the joint distribution of (X_n, S_n) and (T_r, S_{T_r}) for the class of Markov chain embeddable variables of binomial type. In the same spirit, Inoue (2004) established analogous results in the case of multistate trials.

The purpose of this article is to show that, at least in the iid case, the joint distributions of (X_n, S_n), (T_r, S_{T_r}) etc., are uniquely determined by the distribution of X_n or T_r. More specifically in the following section we present results revealing that the probability generating function (pgf) of (X_n, S_n, F_n) (resp. (T_r, S_{T_r}, F_{T_r})) can be directly derived from the pgf of X_n (resp. T_r). Therefore, many of the results concerning joint distributions of this form proved in previous papers using various techniques could be immediately derived. Moreover, we show that the joint pgf of (X_n, S_n, F_n) and (T_r, S_{T_r}, F_{T_r}) are related through a simple formula.

In the final section we present briefly some applications of our main results for some special choices of patterns $\mathcal{E}$.

2. Main results

Let $Z_1, Z_2, ..., Z_n$, $n \geq 1$, be a sequence of iid Bernoulli random variables taking on the values 1 (success, S) or 0 (failure, F) with probabilities $p = P(Z_i = 1)$, $q = P(Z_i = 0) = 1 - p$ ($Z_i \sim Be(p)$, $i = 1, 2, ...$), and denote by $\mathcal{E}$ any pattern, i.e. a string (or a collection of strings) with a prespecified composition. The number of appearances of $\mathcal{E}$ and the number of successes (resp. failures) among $Z_1, Z_2, ..., Z_n$, will be denoted by X_n and S_n (resp. F_n). The probability mass function (pmf) of X_n will be denoted by

$$f_n(x \,|\, p) = \Pr(X_n = x \,|\, Z_i \sim Be(p)), \quad x = 0, 1, ...$$

and the joint pmf of (X_n, S_n) will be denoted by

$$f_n(x, y \,|\, p) = \Pr(X_n = x,\ S_n = y \,|\, Z_i \sim Be(p)), \quad x = 0, 1, ..., \quad y = 0, 1, ..., n.$$

The pgf of X_n and the joint pgf of (X_n, S_n) are given by

$$E(w^{X_n} \,|\, p) = \sum_{x=0}^{\infty} f_n(x \,|\, p) w^x, \qquad E(w^{X_n} w_1^{S_n} \,|\, p) = \sum_{x=0}^{\infty} \sum_{y=0}^{n} f_n(x, y \,|\, p) w^x w_1^y.$$

The following proposition reveals that by applying a simple transformation to the success probability p of the pgf $E(w^{X_n} \,|\, p)$ of X_n we obtain the joint pgf $E(w^{X_n} w_1^{S_n} \,|\, p)$ of (X_n, S_n).

Proposition 2.1. *The joint pgf $E(w^{X_n} w_1^{S_n} \,|\, p)$ of (X_n, S_n) follows from the pgf $E(w^{X_n} \,|\, p)$ of X_n through the relation*

$$E(w^{X_n} w_1^{S_n} \,|\, p) = (pw_1 + q)^n E(w^{X_n} \,|\, \tfrac{pw_1}{pw_1 + q}).$$

Proof. The pmf $f_n(x, y \,|\, p)$ of (X_n, S_n) may be written in the form

$$f_n(x, y \,|\, p) = a_n(x, y) p^y q^{n-y}, \quad x = 0, 1, ..., \quad y = 0, 1, ..., n$$

where $a_n(x, y)$ denotes the number of arrangements of y successes and $n - y$ failures such that in each arrangement the number of appearances of the pattern $\mathcal{E}$ is equal to

x. Exploiting the aforementioned form of $f_n(x,y\,|\,p)$ we have that

$$\begin{aligned} E(w^{X_n}w_1^{S_n}\,|\,p) &= \sum_{x=0}^{\infty}\sum_{y=0}^{n} a_n(x,y)p^y q^{n-y} w^x w_1^y \\ &= (pw_1+q)^n \sum_{x=0}^{\infty}\sum_{y=0}^{n} a_n(x,y)(\tfrac{pw_1}{pw_1+q})^y(\tfrac{q}{pw_1+q})^{n-y} w^x \\ &= (pw_1+q)^n \sum_{x=0}^{\infty}\sum_{y=0}^{n} f_n(x,y\,|\,\tfrac{pw_1}{pw_1+q})w^x \\ &= (pw_1+q)^n \sum_{x=0}^{\infty} f_n(x\,|\,\tfrac{pw_1}{pw_1+q})w^x \\ &= (pw_1+q)^n E(w^{X_n}\,|\,\tfrac{pw_1}{pw_1+q}) \end{aligned}$$

which completes the proof of the proposition. ■

Using Proposition 2.1 we may easily derive the joint pgf $E(w^{X_n}w_1^{S_n}w_0^{F_n}\,|\,p)$ of (X_n,S_n,F_n) in terms of the pgf $E(w^{X_n}\,|\,p)$ of X_n. More specifically, we have that

$$\begin{aligned} E(w^{X_n}w_1^{S_n}w_0^{F_n}\,|\,p) &= E(w^{X_n}w_1^{S_n}w_0^{n-S_n}\,|\,p) = w_0^n E(w^{X_n}\left(\tfrac{w_1}{w_0}\right)^{S_n}\,|\,p) \\ &= (pw_1+qw_0)^n E(w^{X_n}\,|\,\tfrac{pw_1}{pw_1+qw_0}). \end{aligned} \tag{2.1}$$

In some cases only the double generating function of X_n is available. Therefore it would be convenient to obtain a relationship between the double generating functions of (X_n,S_n,F_n) and X_n, which will be denoted by

$$G(w,w_1,w_0;z\,|\,p) = \sum_{n=0}^{\infty} E(w^{X_n}w_1^{S_n}w_0^{F_n}\,|\,p)z^n \quad \text{and} \quad \mathcal{G}(w;z\,|\,p) = \sum_{n=0}^{\infty} E(w^{X_n}\,|\,p)z^n$$

respectively (the coefficients of z^0 in the above two generating functions are conventionally taken to be equal to 1). A straightforward application of (2.1) readily leads to

$$G(w,w_1,w_0;z\,|\,p) = \mathcal{G}(w;z(pw_1+qw_0)\,|\,\tfrac{pw_1}{pw_1+qw_0}). \tag{2.2}$$

Another random variable of interest in the present context is the waiting time T_r, $r \geq 1$, until the r-th appearance of $\mathcal{E}$ in the infinite sequence $Z_1, Z_2, \ldots$. The pmf of T_r will be denoted by

$$h_r(n\,|\,p) = \Pr(T_r = n\,|\,Z_i \sim Be(p)), \quad n = 1,2,\ldots \, .$$

The number of successes (resp. failures) among $Z_1, Z_2, ..., Z_{T_r}$ will be denoted by S_{T_r} (resp. F_{T_r}), and the joint pmf of (T_r, S_{T_r}) by

$$h_r(n, y \,|\, p) = \Pr(T_r = n,\ S_{T_r} = y \,|\, Z_i \sim Be(p)), \quad n = 1, 2, ..., \quad y = 0, 1, ..., n.$$

The pgf of T_r and the joint pgf of (T_r, S_{T_r}) are given by

$$E(z^{T_r} \,|\, p) = \sum_{n=1}^{\infty} h_r(n \,|\, p) z^n, \qquad E(z^{T_r} z_1^{S_{T_r}} \,|\, p) = \sum_{n=1}^{\infty} \sum_{y=0}^{n} h_r(n, y \,|\, p) z^n z_1^y.$$

For the joint pgf of (T_r, S_{T_r}) we may state the following result.

Proposition 2.2. *The joint pgf* $E(z^{T_r} z_1^{S_{T_r}} \,|\, p)$ *of* (T_r, S_{T_r}) *follows from the pgf* $E(z^{T_r} \,|\, p)$ *of* T_r *through the relation*

$$E(z^{T_r} z_1^{S_{T_r}} \,|\, p) = E\left((z(pz_1 + q))^{T_r} \,|\, \tfrac{pz_1}{pz_1+q}\right).$$

Proof. The pmf of (T_r, S_{T_r}) may be written in the form

$$h_r(n, y \,|\, p) = b_r(n, y) p^y q^{n-y}, \quad n = 1, 2, ..., \quad y = 0, 1, ..., n,$$

where $b_r(n, y)$ denotes the number of arrangements of y successes and $n - y$ failures such that at the last element of each arrangement the pattern $\mathcal{E}$ occurs for the rth time. Similar to the proof of Proposition 2.1 we have that

$$\begin{aligned} E(z^{T_r} z_1^{S_{T_r}} \,|\, p) &= \sum_{n=1}^{\infty} \sum_{y=0}^{n} b_r(n, y) p^y q^{n-y} z^n z_1^y \\ &= \sum_{n=1}^{\infty} \sum_{y=0}^{n} b_r(n, y) (\tfrac{pz_1}{pz_1+q})^y (\tfrac{q}{pz_1+q})^{n-y} (z(pz_1 + q))^n \\ &= \sum_{n=1}^{\infty} \sum_{y=0}^{n} h_r(n, y \,|\, \tfrac{pz_1}{pz_1+q})\, (z(pz_1 + q))^n \\ &= \sum_{n=1}^{\infty} h_r(n \,|\, \tfrac{pz_1}{pz_1+q})\, (z(pz_1 + q))^n \\ &= E\left((z(q + pz_1))^{T_r} \,|\, \tfrac{pz_1}{pz_1+q}\right) \end{aligned}$$

which completes the proof of the proposition. ■

Making use of Proposition 2.2 we may easily obtain the joint pgf $E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p)$ of (T_r, S_{T_r}, F_{T_r}) in terms of the pgf $E\left(z^{T_r} \,|\, p\right)$ of T_r, which is given by

$$\begin{aligned} E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p) &= E(z^{T_r} z_1^{S_{T_r}} z_0^{T_r - S_{T_r}} \,|\, p) = E((zz_0)^{T_r} (z_1/z_0)^{S_{T_r}} \,|\, p) \\ &= E\left((z(qz_0 + pz_1))^{T_r} \,|\, \tfrac{pz_1}{pz_1+qz_0}\right). \qquad (2.3) \end{aligned}$$

Next, let

$$H(z, z_1, z_0; w \,|\, p) = \sum_{r=0}^{\infty} E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p) w^r \quad \text{and} \quad \mathcal{H}(z; w \,|\, p) = \sum_{r=0}^{\infty} E(z^{T_r} \,|\, p) w^r$$

be the double generating function of (T_r, S_{T_r}, F_{T_r}) and T_r, respectively (again, the coefficients of w^0 in the above two generating functions are conventionally taken to be equal to 1). Using (2.3) we obtain that $H(z, z_1, z_0; w \,|\, p)$ may be expressed in terms of $\mathcal{H}(z; w \,|\, p)$ as

$$\begin{aligned} H(z, z_1, z_0; w \,|\, p) &= \sum_{r=0}^{\infty} E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p) w^r \\ &= \sum_{r=0}^{\infty} E\left((z(pz_1 + qz_0))^{T_r} \,|\, \tfrac{pz_1}{pz_1 + qz_0} \right) w^r \\ &= \mathcal{H}(z(qz_0 + pz_1); w \,|\, \tfrac{pz_1}{pz_1 + qz_0}). \end{aligned} \tag{2.4}$$

Before we proceed to the derivation of our last result, we note that the pmf of T_r and X_n are related through the obvious identity $\Pr(X_n \geq r) = \Pr(T_r \leq n)$, $n, r \geq 1$, which offers a useful tool for passing from the study of the waiting time for a pattern to the study of the number of appearances of the pattern in fixed number of trials and vice versa. Exploiting this relationship, Koutras (1997) proved, among other results, that the double generating functions $\mathcal{G}(w; z \,|\, p)$ of X_n and $\mathcal{H}(z; w \,|\, p)$ of T_r are related by the formula

$$\mathcal{G}(w; z \,|\, p) = \frac{(w-1)\mathcal{H}(z; w \,|\, p) + 1}{w(1-z)}. \tag{2.5}$$

The above result can now be extended offering the following interrelation between the joint distribution of (X_n, S_n, F_n) and (T_r, S_{T_r}, F_{T_r}).

Proposition 2.3. *The double generating function $G(w, w_1, w_0; z \,|\, p)$ of (X_n, S_n, F_n) can be expressed in terms of the double generating function $H(z, w_1, w_0; w \,|\, p)$ of (T_r, S_{T_r}, F_{T_r}) as*

$$G(w, w_1, w_0; z \,|\, p) = \frac{(w-1)H(z, w_1, w_0; w \,|\, p) + 1}{w(1 - z(pw_1 + qw_0))}.$$

Proof. Using successively relations (2.2), (2.5) and (2.4) we get

$$\begin{aligned} G(w, w_1, w_0; z \,|\, p) &= \mathcal{G}(w; z(qw_0 + pw_1) \,|\, \tfrac{pw_1}{pw_1 + qw_0}) \\ &= \frac{(w-1)\mathcal{H}(z(qw_0 + pw_1); w \,|\, \frac{pw_1}{pw_1 + qw_0}) + 1}{w(1 - z(pw_1 + qw_0))} \\ &= \frac{(w-1)H(z, w_1, w_0; w \,|\, p) + 1}{w(1 - z(pw_1 + qw_0))} \end{aligned}$$

which completes the proof of the proposition. ■

It is worth mentioning that the inversion of the formula established in Proposition 2.3 leads to an expression of $H(z, z_1, z_0; w \,|\, p)$ in terms of $G(w, z_1, z_0; z \,|\, p)$, namely

$$H(z, z_1, z_0; w \,|\, p) = \frac{w(1 - z(qz_0 + pz_1))G(w, z_1, z_0; z \,|\, p) - 1}{w - 1}. \tag{2.6}$$

Finally, we note that all the power series of the form $\sum_{i_1,\ldots,i_k} a_{i_1,\ldots,i_k} w_1^{i_1} \ldots w_k^{i_k}$ we employ are convergent with respect to all their parameters $w_1, \ldots, w_k$ *at least* in a neighborhood of zero of the form $(0, r_1) \times \cdots \times (0, r_k)$. By taking $w_1 \to 0^+, \ldots, w_k \to 0^+$ in appropriate partial derivatives of the power series (which exist in the same neighborhood) we can recover all the $a_{i_1,\ldots,i_k}$ and thus all pgfs we employ uniquely determine their underlying distribution. We restrict ourselves in positive $w_1, \ldots, w_k$ and not, as usually, in a neighborhood of the form $(-r_1, r_1) \times \cdots \times (-r_k, r_k)$, in order to avoid possible appearances of negative probabilities in expressions like $E(w^{X_n} \mid \frac{pw_1}{pw_1+q})$ etc.

3. Applications

In this section we present two interesting applications of the general results established in the previous section. More specifically, we study the appearances of some specific patterns $\mathcal{E}$ in a sequence of Bernoulli trials (with success probability p) illustrating how one can instantly obtain pgfs of joint distributions of successes, failures and patterns relying only on the pgf of T_r or X_n.

Example 1. Let T_r be the waiting time for the r-th appearance of the composite pattern

$$\mathcal{E} = \{11, 101, 1001, \ldots, 1\underbrace{00\ldots0}_{k-2}1\}$$

($k \geq 2$). Antzoulakos (2001) by employing the overlapping counting scheme derived that

$$E(z^{T_r} \,|\, p) = \frac{(pz)^2(1 - (qz)^{k-1})}{(1 - qz)(1 - qz - (pz)(qz)^{k-1})} \left(\frac{(pz)(1 - (qz)^{k-1})}{1 - qz - (pz)(qz)^{k-1}} \right)^{r-1}, \quad r \geq 1.$$

Using relation (2.3) we immediately obtain that

$$E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p) = \frac{pzz_1}{1 - qzz_0} \left(\frac{(pzz_1)(1 - (qzz_0)^{k-1})}{1 - qzz_0 - (pzz_1)(qzz_0)^{k-1}} \right)^r, \quad r \geq 1. \tag{2.7}$$

The pgfs of S_{T_r}, F_{T_r}, $(T_r,\, S_{T_r})$, $(T_r,\, F_{T_r})$ and $(S_{T_r},\, F_{T_r})$ follow readily from the above formula, which are all new results. It is noteworthy that S_{T_r} follows a negative binomial distribution with parameters r and $1-q^{k-1}$ shifted to the support $\{r+1, r+2, ...\}$.

Furthermore, making use of relation (2.7) we may derive the double generating function of (T_r, S_{T_r}, F_{T_r}) which is given by

$$H(z, z_1, z_0; w \,|\, p) = 1 + \frac{w(pzz_1)^2(1-(qzz_0)^{k-1})}{(1-qzz_0)(1-qzz_0-w(pzz_1)-(pzz_1)(qzz_0)^{k-1}(1-w))}.$$

Using Proposition 2.3 we obtain that

$$G(w, w_1, w_0; z \,|\, p) = \frac{1-qzw_0+pzw_1-w(pzw_1)-(pzw_1)(qzw_0)^{k-1}(1-w)}{(1-qzw_0)(1-qzw_0-w(pzw_1)-(pzw_1)(qzw_0)^{k-1}(1-w))}.$$

The above expression can be exploited to obtain the following recursive scheme

$$\begin{aligned}
G_0(\mathbf{w} \,|\, p) &= 1, \quad G_1(\mathbf{w} \,|\, p) = pw_1 + qw_0, \quad G_2(\mathbf{w} \,|\, p) = (qw_0)^2 + 2(qw_0)(pw_1) + w(pw_1)^2 \\
G_n(\mathbf{w} \,|\, p) &= (2qw_0 + wpw_1)G_{n-1}(\mathbf{w} \,|\, p) - (qw_0)(qw_0 + wpw_1)G_{n-2}(\mathbf{w} \,|\, p), \qquad 3 \le n \le k \\
G_n(\mathbf{w} \,|\, p) &= (2qw_0 + wpw_1)G_{n-1}(\mathbf{w} \,|\, p) - (qw_0)(qw_0 + wpw_1)G_{n-2}(\mathbf{w} \,|\, p) \\
&\quad + (1-w)(pw_1)(qw_0)^{k-1}[G_{n-k}(\mathbf{w} \,|\, p) - (qw_0)G_{n-k-1}(\mathbf{w} \,|\, p)], \qquad n \ge k+1
\end{aligned}$$

satisfied by the joint pgf

$$G_n(\mathbf{w} \,|\, p) = G_n(w, w_1, w_0 \,|\, p) = E(w^{X_n} w_1^{S_n} w_0^{F_n} \,|\, p)$$

of (X_n, S_n, F_n). Further manipulation over the above recursion may provide a recursive scheme satisfied by the joint pmf of (X_n, S_n, F_n).

In closing Example 1, we note that Koutras (1996) derived the pgf of T_r in the case of the non-overlapping way of counting runs and patterns which is given by

$$E(z^{T_r} \,|\, p) = \left(\frac{(pz)^2(1-(qz)^{k-1})}{(1-qz)(1-qz-(pz)(qz)^{k-1})} \right)^r, \quad r \ge 1.$$

Using relation (2.3) we get that

$$E(z^{T_r} z_1^{S_{T_r}} z_0^{F_{T_r}} \,|\, p) = \left(\frac{(pzz_1)^2(1-(qzz_0)^{k-1})}{(1-qzz_0)(1-qzz_0-(pzz_1)(qzz_0)^{k-1})} \right)^r, \quad r \ge 1,$$

from which several results established by Chadjiconstantinidis and Koutras (2001) may be recovered. ■

Example 2. Let X_n denote the number of success runs of length at least k in n iid Bernoulli trials with success probability p. The double generating function of X_n is given by

$$\mathcal{G}(w;z\,|\,p)=\frac{1-(1-w)(pz)^k}{1-z+(1-w)qp^kz^{k+1}}$$

(see e.g. Hirano and Aki (1993) and Koutras and Alexandrou (1995)). A straightforward application of relation (2.2) yields

$$G(w,w_1,w_0;z\,|\,p)=\frac{1-(1-w)(pzw_1)^k}{1-z(pw_1+qw_0)+(1-w)(qzw_0)(pzw_1)^k}.$$

The above expression implies the following recursive scheme satisfied by the joint pgf $G_n(\mathbf{w}\,|\,p)$ of (X_n,S_n,F_n)

$$\begin{aligned}
G_0(\mathbf{w}\,|\,p) &= 1,\quad G_n(\mathbf{w}\,|\,p)=(qw_0+pw_1)^n, \qquad 1\le n\le k-1\\
G_k(\mathbf{w}\,|\,p) &= (qw_0+pw_1)^k-(1-w)(pw_1)^k\\
G_n(\mathbf{w}\,|\,p) &= (qw_0+pw_1)G_{n-1}(\mathbf{w}\,|\,p)-(1-w)(qw_0)(pw_1)^kG_{n-k-1}(\mathbf{w}\,|\,p),\quad n\ge k+1.
\end{aligned}$$

For $w_0=w_1=1$ the above recursion coincides with the respective one obtained by Antzoulakos and Chadjiconsantinidis (2001).

The application of relation (2.6) leads to the following expression for the double generating function of (T_r,S_{T_r},F_{T_r})

$$\begin{aligned}
H(z,z_1,z_0;w\,|\,p) &= \frac{1-z(pz_1+qz_0)+(1-w)(qzz_0)(pzz_1)^k+w(pzz_1)^k(1-pzz_1)}{1-z(pz_1+qz_0)+(1-w)(qzz_0)(pzz_1)^k}\\
&= \frac{1-w\left(\frac{qzz_0}{1-pzz_1}-1\right)Q(z,z_1,z_0\,|\,p)}{1-w\left(\frac{qzz_0}{1-pzz_1}\right)Q(z,z_1,z_0\,|\,p)}
\end{aligned}$$

where

$$Q(z,z_1,z_0\,|\,p)=\frac{(pzz_1)^k(1-pzz_1)}{1-z(qz_0+pz_1)+(qzz_0)(pzz_1)^k}.$$

Expanding $H(z,z_1,z_0;w\,|\,p)$ in powers of w we obtain that

$$E(z^{T_r}z_1^{S_{T_r}}z_0^{F_{T_r}}\,|\,p)=Q(z,z_1,z_0\,|\,p)\left(\tfrac{qzz_0}{1-pzz_1}Q(z,z_1,z_0\,|\,p)\right)^{r-1},\quad r\ge 1.$$

We note that some special results of the present example has also been obtained by Chadjiconstantinidis *et al.* (2000) via a different methodology. ■

In closing we mention that even though we presented a methodology to obtain the joint pgf of (T_r,S_{T_r},F_{T_r}) and (X_n,S_n,F_n) through the distribution of X_n or T_r, the derivation of closed formulae or recursive schemes for the pmfs of the desired distributions seems to be a difficult task in cases where these pgfs are very complicated. Nevertheless, we can always obtain the pmf from a pgf (e.g. by repeated differentiation of the pgf) relying on the growing computational power offered by modern computer algebra software (e.g. Mathematica, Matlab, etc.).

References

Aki, S. and Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes, *Ann. Inst. Statist. Math.,* **46**, 193-202.

Aki, S. and Hirano, K. (1995). Joint distributions of numbers of success-runs and failures until the first consecutive k successes, *Ann. Inst. Statist. Math.*, **47**, 225-235.

Antzoulakos, D. L. (2001). Waiting times for patterns in a sequence of multistate trials, *J. Appl. Prob.*, **38**, 508-518.

Antzoulakos, D. L. and Chadjiconstantinidis, S. (2001). Distributions of numbers of success runs of fixed length in Markov dependent trials. *Ann. Inst. Statist. Math.* **53**, 599-601.

Antzoulakos, D. L. and Philippou, A. N. (1996). Distributions of the numbers of successes and failures until the occurrence of the r-th overlapping and non-overlapping success run of length k, *Journal of the Indian Society of Statistics and Operations Research*, **17**, 19-30.

Balakrishnan, N. (1997). Joint distributions of numbers of success-runs and failures until the first consecutive k successes in a binary sequence, *Ann. Inst. Statist. Math.*, **49**, 519-529.

Balakrishnan, N. and Koutras, M. V. (2002). *Runs, Scans and Applications.*, John Wiley & Sons.

Chadjiconstantinidis, S. and Koutras, M. V. (2001). Distributions of the numbers of failures and successes in a waiting time problem, *Ann. Inst. Statist. Math.,* **53**, 576-598.

Chadjiconstantinidis, S., Antzoulakos, D. L. and Koutras, M. V. (2000). Joint distributions of successes, failures and patterns in enumeration problems, *Adv. Appl. Prob.*, **32**, 866-884.

Fu, J. C. and Lou, W. Y. W. (2003). *Distribution Theory of Runs and Patterns and Its Applications*, World Scientific, New Jersey.

Glaz, J. and Balakrishnan, N. (1999).Introduction to scan statistics. In *Scan Statistics and Applications* (Eds., J. Glaz and N. Balakrishnan), 3-24, Birkhäuser, Boston.

Glaz, J. , Naus, J. and Wallenstein, S.(2001). *Scan Statistics.* Springer.

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

Inoue, K. (2004). Joint distributions associated with patterns, successes and failures in a sequence of multi-state trials, *Ann. Inst. Statist. Math.,* **56**, 143-168.

Koutras, M. V. (1996). On a waiting time distribution in a sequence of Bernoulli trials, *Ann. Inst. Statist. Math.*, **48**, 789-806.

Koutras, M. V. (1997). Waiting times and number of appearances of events in a sequence of discrete random variables, *Advances in Combinatorial Methods and Applications to Probability and Statistics*, (ed. N. Balakrishnan), 363-384, Birkhauser, Boston.

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.

Uchida, M. (1998). Joint distributions of the numbers of success runs until the first consecutive k successes in a higher-order two-state Markov chain. *Ann. Inst. Statist. Math.*, **50**, 203-222.