RT-MAE 9905

QUOTAS ON RUNS OF SUCCESSES AND FAILURES IN A MULTI-STATE MARKOV CHAIN

by

Nikolai Kolev
and
Leda Minkova

Palavras-Chave: joint probability generating function; sooner and later waiting time problems; time homogeneous multi-state Markov chain; recursion; run quota.

Classificação AMS: 60E05, 60J10.
(AMS Classification)

- Fevereiro de 1999 -

QUOTAS ON RUNS OF SUCCESSES AND FAILURES IN A MULTI-STATE MARKOV CHAIN

Nikolai Kolev[1] **and Leda Minkova**[2]

Department of Statistics - IME
São Paulo University
C.P. 66281, 05315-970 São Paulo, Brazil
([1]Institute of Mathematics and Informatics, Bulg. Acad. Sci.
and [2]Technical University of Sofia, Bulgaria)
e-mail: nkolev@ime.usp.br

Key Words and Phrases: joint probability generating function; sooner and later waiting time problems; time homogeneous multi-state Markov chain; recursion; run quota.

ABSTRACT

For the time-homogeneous multi-state Markov chain $\{X_n, n \geq 0\}$ with states labeled as "0" (success) and "f" (failure), $f = 1, 2, \ldots$ the waiting time problems to be discussed arise by setting quotas on runs of successes and failures. Some particular cases are considered.

1. INTRODUCTION

We consider the time-homogeneous multi-state Markov chain $\{X_n, n \geq 0\}$ taking values $\{0, 1, 2, \ldots\}$, where the sequence $\{X_0, X_1, \ldots\}$ is determined by the distribution of the initial states

$$P(X_0 = g_1) = p_{g_1},\ 0 < p_0 < 1,\ p_{g_1} \geq 0,\ g_1 = 1, 2, \ldots,\ \sum_{g_1=0}^{\infty} p_{g_1} = 1,$$

and transition probabilities

$$P(X_{n+1} = g_2 \mid X_n = g_1) = p_{g_1 g_2}, \quad g_1, g_2 = 0, 1, \ldots, \quad n = 0, 1, \ldots$$

where

$$0 < p_{00} < 1, \quad \sum_{g_2=0}^{\infty} p_{g_1 g_2} = 1, \quad g_1 = 0, 1, \ldots$$

and

$$p_{g_1 g_2} \geq 0, \quad g_1 = 1, 2, \ldots, \quad g_2 = 0, 1, \ldots.$$

We regard the value "0" as success and the remaining values "f" as failures, $f = 1, 2, \ldots$.

The waiting times to be discussed will arise by setting quotas on runs of successes and failures.

Let c and d be fixed positive integers and let us consider the events

$$\mathbf{ST}(c,d) = \left\{ \begin{array}{c} \text{the trials are performed sequentially} \\ \text{until either } c \text{ consecutive successes} \\ \text{or } d \text{ consecutive failures are observed} \end{array} \right\},$$

$$\mathbf{ST}_1(c) = \left\{ \begin{array}{c} \text{the trials are performed sequentially} \\ \text{until } c \text{ consecutive successes are observed} \end{array} \right\},$$

$$\mathbf{ST}_2(d) = \left\{ \begin{array}{c} \text{the trials are performed sequentially} \\ \text{until } d \text{ consecutive failures are observed} \end{array} \right\}$$

and

$$\mathbf{LT}(c,d) = \left\{ \begin{array}{c} \text{the trials are performed sequentially upon} \\ \text{completion of both a run of } c \text{ successes} \\ \text{and a run of } d \text{ failures} \end{array} \right\}.$$

It is clear that the sooner occurring event $\mathbf{ST}(c,d)$ and the later occurring event $\mathbf{LT}(c,d)$ between $\mathbf{ST}_1(c)$ and $\mathbf{ST}_2(d)$ are determined by the relations

$$\mathbf{ST}(c,d) = \mathbf{ST}_1(c) \cup \mathbf{ST}_2(d) \tag{1}$$

and

$$\mathbf{LT}(c,d) = \mathbf{ST}_1(c) \cap \mathbf{ST}_2(d). \tag{2}$$

In this paper we extend to the present homogeneous Markov chain the waiting time problems related with the events $\mathbf{ST}(c,d)$ and $\mathbf{LT}(c,d)$. These events are studied by Balasubramanian et *al.* (1993) for Markov correlated Bernoulli trials, where the review of literature is given.

Similar results for two-state homogeneous Markov chain are obtained by Aki and Hirano (1993) and Uchida and Aki (1995). The distribution theory of runs in different Markov fashions is developed by many authors. Some important recent references are Aki et *al.* (1996), Balakrishnan et *al.* (1997), Doi and Yamamoto (1998), Ebneshahrashoob and Sobel (1995), Fu and Kuotras (1996), Kuotras (1997).

Let us define the random variables

$$S_{g_1} = \begin{cases} 1 & \text{if } X_0 = g_1, \\ 0 & \text{if } X_0 \neq g_1, \end{cases}$$

for $g_1 = 0, 1, \ldots$ and

$$S_{g_1 g_2} = \left\{ \begin{array}{c} \text{number of transitions of type } g_1 \longrightarrow g_2 \text{ in the} \\ \text{Markov chain until the event } \mathbf{ST}(c,d) \text{ occurs} \end{array} \right\},$$

for $g_1, g_2 = 0, 1, \ldots$.

In Section 2 we deal with the distribution of the random vector

$$\mathbf{S} = (S_0, S_1, \ldots, S_{00}, S_{01}, \ldots),$$

under the condition that the event $\mathbf{ST}(c,d)$ occurs with probability 1. To avoid trivial complications we assume the chain to be irreducible. As particular cases, the exact joint distribution of the number of successes and number of failures is given as well as a run quota on successes and a run quota of failures are studied. Two examples are given too. Because of duality based on the relations (1) and (2) the later cases related with the event $\mathbf{LT}(c,d)$ are derived from Section 2 and are discussed in Section 3. Some final remarks are given in the last section.

2. SOONER WAITING TIME PROBLEMS

Under the given notations the following basic lemma is true.

Lemma 2.1. *The joint probability generating function (PGF) of the random vector* $\mathbf{S} = (S_0, S_1, \ldots, S_{00}, S_{01}, \ldots)$ *related with the event* $\mathbf{ST}(c,d)$ *is given by*

$$\begin{aligned} g_{\mathbf{S}}(t_0, t_1, \ldots, t_{00}, t_{01}, \ldots) &= E\{(t_0)^{S_0}(t_1)^{S_1} \ldots (t_{00})^{S_{00}}(t_{01})^{S_{01}} \ldots\} \\ &= \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1}^{(d)} + \frac{A_0 C_{00}}{1 - A_{00}} \end{aligned} \qquad (3)$$

which converges at least for

$$\{(t_0, t_1, \ldots, t_{00}, t_{01}, \ldots) : |t_i| \le 1 \ \ \textit{and} \ \ |t_{ij}| \le 1, \ \ i,j = 0,1,\ldots\}$$

and

$$A_0 = p_0 t_0 + \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1 0}^{(d-1)}, \tag{4}$$

$$A_{00} = \frac{1-(p_{00}t_{00})^{c-1}}{1-p_{00}t_{00}} \sum_{g_1=1}^{\infty} p_{0g_1} t_{0g_1} F_{g_1 0}^{(d-1)}, \tag{5}$$

$$C_{00} = (p_{00}t_{00})^{c-1} + \frac{1-(p_{00}t_{00})^{c-1}}{1-p_{00}t_{00}} \sum_{g_1=1}^{\infty} p_{0g_1} t_{0g_1} F_{g_1}^{(d)},$$

where for $g_1 = 1, 2, \ldots$ *and* $d > 1$

$$F_{g_1}^{(d)} = \sum_{g_2=1}^{\infty} p_{g_1g_2} t_{g_1g_2} \cdots \sum_{g_d=1}^{\infty} p_{g_{d-1}g_d} t_{g_{d-1}g_d}, \tag{6}$$

$$\begin{aligned} F_{g_1 0}^{(d-1)} &= p_{g_1 0} t_{g_1 0} + \sum_{g_2=1}^{\infty} p_{g_1g_2} t_{g_1g_2} p_{g_2 0} t_{g_2 0} + \ldots \\ &+ \sum_{g_2=1}^{\infty} p_{g_1g_2} t_{g_1g_2} \cdots \sum_{g_{d-1}=1}^{\infty} p_{g_{d-2}g_{d-1}} t_{g_{d-2}g_{d-1}} p_{g_{d-1}0} t_{g_{d-1}0}. \end{aligned} \tag{7}$$

If $d = 1$ *the expressions (6) and (7) take the form*

$$F_{g_1}^{(d)} = 1 \quad \textit{and} \quad F_{g_1 0}^{(d-1)} = 0.$$

Proof. Let us consider at first the sequences of outcomes terminating in c consecutive successes "0". The possible disjoint events for $d > 1$ are as follows:

– $X_0 = 0$ and the first success is followed by $c-1$ consecutive successes. The contribution to the joint PGF $g_S(\bullet)$ is

$$S^0 = p_0 t_0 (p_{00}t_{00})^{c-1};$$

– $X_0 = 0$ and exactly $r \ge 1$ subsequences containing no more than $d-1$ failures are observed.

The possible sequences of outcomes can be described as

$$0\underbrace{00\ldots0}_{i_1}\underbrace{ff\ldots f}_{j_1}\underbrace{00\ldots0}_{i_2}\underbrace{ff\ldots f}_{j_2}\ldots\underbrace{00\ldots0}_{i_r}\underbrace{ff\ldots f}_{j_r}\underbrace{00\ldots0}_{c}$$

where

$$0 \le i_1 \le c-2; \ 1 \le i_a \le c-1, \ a = 2,3,\ldots,r; \ 1 \le j_b < d, \ b = 1,2,\ldots,r.$$

From the Markov property of the sequences follows that the net contribution S_r^1 to the joint PGF $g_S(\bullet)$ of all such sequences is

$$S_r^1 = p_0 t_0 (A_{00})^r (p_{00} t_{00})^{c-1},$$

where A_{00} is given by (5);

- $X_0 = f$, $f = 1, 2, \ldots$ and exactly $r \geq 1$ subsequences containing no more than $d-1$ failures are observed.

The contribution to the joint PGF in this case is

$$S_r^2 = \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1 0}^{(d-1)} (A_{00})^{r-1} (p_{00} t_{00})^{c-1}.$$

Adding of all of these contributions gives

$$S = S^0 + \sum_{r=1}^{\infty} (S_r^1 + S_r^2)$$

and after simplifications we obtain

$$S = \frac{A_0}{1 - A_{00}} (p_{00} t_{00})^{c-1}.$$

Let us consider the sequences of outcomes terminating in d consecutive failures $f = 1, 2, \ldots$ when $d > 1$.

Let $X_0 = f$ and the first failure is followed by $d-1$ consecutive failures. The contribution to the joint PGF in this case is

$$F_1 = \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1}^{(d)}.$$

Next, the contribution to the joint PGF, when the initial trial is successful followed by no more than $c-1$ consecutive successes preceded the sequence of d consecutive failures, is given by

$$F_2 = p_0 t_0 \frac{1 - (p_{00} t_{00})^{c-1}}{1 - p_{00} t_{00}} \sum_{g_1=1}^{\infty} p_{0 g_1} t_{0 g_1} F_{g_1}^{(d)}.$$

The corresponding contributions F_r^1 and F_r^2 from the sequences ending with d consecutive failures, containing exactly $r \geq 1$ subsequences of failures and beginning with a success or a failure are

$$F_r^1 = p_0 t_0 (A_{00})^r \frac{1 - (p_{00} t_{00})^{c-1}}{1 - p_{00} t_{00}} \sum_{g_1=1}^{\infty} p_{0 g_1} t_{0 g_1} F_{g_1}^{(d)}$$

and

$$F_r^2 = \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1 0}^{(d-1)} (A_{00})^{r-1} \frac{1-(p_{00}t_{00})^{c-1}}{1-p_{00}t_{00}} \sum_{g_1=1}^{\infty} p_{0g_1} t_{0g_1} F_{g_1}^{(d)},$$

where $A_{00}, F_{g_1}^{(d)}$ and $F_{g_1 0}^{(d-1)}$ are given by (5), (6) and (7).

Adding of all of these contributions gives

$$F = F_1 + F_2 + \sum_{r=1}^{\infty}(F_r^1 + F_r^2),$$

i.e.

$$F = \sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} F_{g_1}^{(d)} + \frac{A_0}{1-A_{00}} \left[C_{00} - (p_{00}t_{00})^{c-1}\right].$$

The joint PGF $g_{\mathbf{S}}(\bullet)$ given by (3) is obtained by

$$g_{\mathbf{S}}(t_0, t_1, \ldots, t_{00}, t_{01}, \ldots) = S + F.$$

At the end, if $d = 1$, the joint PGF is given by the following expression

$$\sum_{g_1=1}^{\infty} p_{g_1} t_{g_1} + p_0 t_0 \left[(p_{00}t_{00})^{c-1} + \frac{1-(p_{00}t_{00})^{c-1}}{1-p_{00}t_{00}} \sum_{g_1=1}^{\infty} p_{0g_1} t_{0g_1}\right],$$

i.e. one can substitute $F_{g_1}^{(d)} = 1$ and $F_{g_1 0}^{(d-1)} = 0$ in (3).

To see that $g_{\mathbf{S}}(\bullet)$ is a PGF, let us put in (3) all arguments equal to 1 and $\mathbf{1} = (1, 1, \ldots, 1, 1, \ldots)$ be the corresponding unit vector. Then the following relations

$$F_{g_1 0}^{(d-1)}(\mathbf{1}) = \sum_{i=1}^{d-1} f_{g_1 0}(i) = 1 - F_{g_1}^{(d)}(\mathbf{1})$$

and

$$1 - A_{00}(\mathbf{1}) = C_{00}(\mathbf{1}),$$

are true, where $f_{g_1 0}(i)$ means the conditional probability that first visit to state 0 from some failure state occurs in exactly i steps, $i = 1, 2, \ldots, d-1$. This completes the proof.

Remark 2.1. Lemma 2.1 is true under condition that $P(\mathbf{ST}(c,d)) = 1$, which is always fulfilled for irreducible Markov chains with persistent set of states.

If the states of the chain are transient, for the realization of the event $\mathbf{ST}(c,d)$ with probability 1, it is necessary all possible sequences of outcomes between the first and the last observed successive state "0" in the event $\mathbf{ST}(c,d)$ to belong to the same state, say $\mathbf{P}_0$, of transient states.

Let $X_0 = 0$ and the sequences of outcomes terminate in d consecutive transient failure states. Beginning from some of the last d failure states it is possible the elements of the following subsequence to belong to the upper (according $\mathbf{P}_0$) sets of states.

Let the chain starts with a failure state. Then some subsequence of consecutive failure states, followed the initial state of the chain and preceding the first successive state, may belong to the lower (according $\mathbf{P}_0$) set of transient states.

Remark 2.2. If the homogeneous Markov chain has two states from (3) one can obtain the corresponding joint PGF given by Balasubramanian et *al.* (1993) in their Proposition 1. Lemma 2.1 generalizes the PGF given in Theorem 1 by Aki and Hirano (1993), where it is understood that X_0 is not counted in measuring the length of the run in two-state Markov chain.

Remark 2.3. By substituting in (3)

$$t_0 = t_{g_1 0} = s, \quad g_1 = 0, 1, \ldots$$

and

$$t_{g_1} = t, \; g_1 = 1, 2, \ldots; \quad t_{g_1 g_2} = t, \;\; g_1 = 0, 1, \ldots \quad \text{and} \quad g_2 = 1, 2, \ldots$$

we obtain the joint PGF

$$g_{S,F}(s,t) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} P(X = x, F = y) s^x t^y \tag{8}$$

of the random vector (S, F), where S and F mean number of successes and failures in the event $\mathbf{ST}(c,d)$, respectively. From (8) it is possible to calculate $Cov(S, F)$.

Let us denote by

$$a_1(j) = \sum_{g_1=1}^{\infty} p_{g_1} f_{g_1 0}(j), \quad j = 1, 2, \ldots, d-1,$$

$$a_0(j) = \sum_{g_1=1}^{\infty} p_{0 g_1} f_{g_1 0}(j), \quad j = 1, 2, \ldots, d-1,$$

$$a_1(d) = \sum_{g_1=1}^{\infty} p_{g_1} F_{g_1}^{(d)}(\mathbf{1})$$

and

$$a_0(d) = \sum_{g_1=1}^{\infty} p_{0 g_1} F_{g_1}^{(d)}(\mathbf{1}),$$

where $f_{g_1 0}(j)$ are the conditional probabilities defined in the proof of Lemma 2.1 and $\mathbf{1}$ means the corresponding unit vector.

From (8) one can obtain the following exact join distribution of the random vector (S, F):

$$\begin{array}{lcll}
P(S=x,F=y) & = & 0 & \text{for} \quad x<c \quad \text{and} \quad y<d; \\
P(S=c,F=0) & = & p_0(p_{00})^{c-1}; & \\
P(S=0,F=d) & = & a_1(d); & \\
P(S=c,F=d) & = & 0; & \\
P(S=c+i,F=0) & = & 0 & \text{for} \quad i=1,2,\ldots; \\
P(S=0,F=d+j) & = & 0 & \text{for} \quad j=1,2,\ldots; \\
P(S=c,F=j) & = & a_1(j)(p_{00})^{c-1} & \text{for} \quad j=1,2,\ldots,d-1; \\
P(S=i,F=d) & = & p_0(p_{00})^{i-1}a_0(d) & \text{for} \quad i=1,2,\ldots,c-1;
\end{array}$$

$$\begin{aligned}
P(S=i,F=d+j) &= (p_{00})^{i-1}[a_1(j)a_0(d)-a_0(j)a_1(d)] \\
&+ \sum_{m=1}^{i}\sum_{n=1}^{j}(p_{00})^{m-1}a_0(n)P(S=i-m,F=d+j-n) \\
&\text{for} \quad i=1,2,\ldots,c-1 \quad \text{and} \quad j=1,2,\ldots,d-1.
\end{aligned}$$

The remaining joint probabilities for $x=c,\ y=d+1,d+2,\ldots,2d-1$ and $x\geq c+1,\ y\geq 2d$, are given recursively by the following relation

$$P(S=x,F=y)=\sum_{m=1}^{c-1}\sum_{n=1}^{d-1}(p_{00})^{m-1}a_0(n)P(S=x-m,F=y-n).$$

From (8) can be obtained the marginal PGF's $g_S(s)$, $g_F(t)$ and $g_N(u)$ of the total number of successes S, the total number of failures F and the total number of trials N in the event $\mathbf{ST}(c,d)$, respectively, as well as the corresponding probability mass functions and expected waiting times.

2.1. EXACT DISTRIBUTION OF THE NUMBER OF SUCCESSES

By substituting in (3)

$$t_0=t_{g_10}=s, \quad g_1=0,1,\ldots$$

and

$$t_{g_1}=1,\ g_1=1,2,\ldots; \quad t_{g_1g_2}=1,\ g_1=0,1,\ldots \quad \text{and} \quad g_2=1,2,\ldots$$

we obtain the PGF of the number of successes S in the event $\mathbf{ST}(c,d)$. The corresponding expression is given by

$$g_S(s)=a_1(d)+\frac{\left[p_0+\sum_{j=1}^{d}a_1(j)\right]\left[(p_{00})^{c-1}s^c+sa_0(d)\frac{1-(p_{00}s)^{c-1}}{1-p_{00}s}\right]}{1-\sum_{m=1}^{c-1}(p_{00})^{m-1}s^m\sum_{j=1}^{d-1}a_0(j)}, \tag{9}$$

where $a_0(d), a_1(d), a_0(j)$ and $a_1(j)$ are as defined in Remark 2.3. Now, for the exact distribution of the number of successes we obtain the following formulas:

$$\begin{aligned} P(S=0) &= a_1(d); \\ P(S=i) &= \left[p_0 + \sum_{j=1}^{d-1} a_1(j)\right] (p_{00})^{i-1} a_0(d) \\ &+ \sum_{m=1}^{i-1} P(S=i-m)(p_{00})^{m-1} \sum_{j=1}^{d-1} a_0(j) \quad \text{for} \quad i=1,2,\ldots,c-1; \\ P(S=c) &= \left[p_0 + \sum_{j=1}^{d-1} a_1(j)\right] (p_{00})^{c-1}. \end{aligned}$$

The remaining probabilities can be calculated recursively

$$P(S=x) = \sum_{m=1}^{c-1} \sum_{j=1}^{d-1} (p_{00})^{m-1} a_0(j) P(S=x-m) \quad \text{for} \quad x > c.$$

Example 2.1. Let us consider a Markov chain having 3 states: a successive one and two types of failure states (1 and 2), with transition probabilities given by the following matrix

$$\begin{pmatrix} p_{00} & p_{01} & 0 \\ p_{10} & 0 & p_{12} \\ p_{20} & 0 & p_{22} \end{pmatrix}$$

and $p_0 + p_1 = 1$.

This matrix corresponds to the single corrective action model for a start-up demonstration test under the Markovian dependence structure studied by Balakrishnan et *al.* (1997). We slightly simplify the model accepting the same probability

$$p_{00} = P(\text{success} \mid \text{previous trial is successive}).$$

In a single corrective action model, one corrective action of the equipment is allowed (failure of type 1), resulting in a change in the probability of success just after a failure, until specified number of consecutive start-ups, say c, are observed. This means that just after the first failure the system changes his behavior. So, it is impossible the test to begin with a failure of the second type. Because of this $p_2 = 0$.

Under these assumptions, one can obtain when $d \to \infty$ from (9) the PGF given in Theorem 2.1 by Balakrishnan et *al.* (1997).

Let us substitute $d = 3$. This means that under the corrective action model, the start-ups are terminated upon the c-th consecutive successive state of the equipment or just after the third observed consecutive failure.

In this case, we have the following simple expressions

$$\begin{aligned}
a_0(1) &= p_{01}p_{10};\\
a_0(2) &= p_{01}p_{12}p_{20};\\
a_0(3) &= p_{01}p_{12}p_{22};\\
a_1(1) &= p_1p_{10};\\
a_1(2) &= p_1p_{12}p_{20};\\
a_1(3) &= p_1p_{12}p_{22},
\end{aligned}$$

which are used in calculating the probabilities $P(S=x),\ x \geq 0$.

Remark 2.4. Letting $d \to \infty$ in the event $\mathbf{ST}(c,d)$ yields the event $\mathbf{ST}_1(c)$ studied by Kolev and Minkova (1997). Then from (3) we obtain the joint PGF $g(c,\infty)$ until a run of c consecutive successes in a multi-state Markov chain is observed, given by

$$g(c,\infty) = \frac{A_0^1}{1-A_{00}^1}(p_{00}t_{00})^{c-1}, \tag{10}$$

where

$$A_0^1 = p_0t_0 + \sum_{g_1=1}^{\infty} p_{g_1}t_{g_1}F_{g_10},$$

$$A_{00}^1 = \frac{1-(p_{00}t_{00})^{c-1}}{1-p_{00}t_{00}} \sum_{g_1=1}^{\infty} p_{0g_1}t_{0g_1}F_{g_10},$$

with

$$F_{g_10} = p_{g_10}t_{g_10} + \sum_{g_2=1}^{\infty} p_{g_1g_2}t_{g_1g_2}p_{g_20}t_{g_20} + \sum_{g_2=1}^{\infty} p_{g_1g_2}t_{g_1g_2} \sum_{g_3=1}^{\infty} p_{g_2g_3}t_{g_2g_3}p_{g_30}t_{g_30} + \cdots$$

Let us note that in this limiting case the relation (10) is true only for irreducible Markov chains with persistent set of states (compare with Remark 2.1).

The expression (10) is preferable to be used since it is simpler than the corresponding joint PGF derived by Kolev and Minkova (1997).

2.2. EXACT DISTRIBUTION OF THE NUMBER OF FAILURES

By analogy with the previous subsection, let us substitute

$$t_0 = t_{g_10} = 1, \quad g_1 = 0, 1, \ldots$$

and

$$t_{g_1} = t,\ g_1 = 1, 2, \ldots; \quad t_{g_1g_2} = t,\ g_1 = 0, 1, \ldots \quad \text{and} \quad g_2 = 1, 2, \ldots$$

As a result, from (3) one can obtain the PGF $g_F(t)$ of the total number F of failures in the event $\mathbf{ST}(c,d)$

$$g_F(t) = a_1(d)t^d + \frac{\left[p_0 + \sum_{j=1}^{d-1} a_1(j)t^j\right]\left[(p_{00})^{c-1} + a_0(d)\frac{1-(p_{00})^{c-1}}{1-p_{00}}t^d\right]}{1 - \frac{1-(p_{00})^{c-1}}{1-p_{00}}\sum_{j=1}^{d-1} a_0(j)t^j},$$

where the expressions for $a_0(d), a_1(d), a_0(j)$ and $a_1(j)$ are given in Remark 2.3. The probability mass function of the total number of failures is given by the following relations

$$\begin{aligned}
P(F=0) &= p_0(p_{00})^{c-1};\\
P(F=j) &= \left[a_1(j) + p_0\frac{1-(p_{00})^{c-1}}{1-p_{00}}a_0(j)\right](p_{00})^{c-1} \quad \text{for } j = 1,2,\ldots,d-1;\\
P(F=d) &= a_1(d) + p_0\frac{1-(p_{00})^{c-1}}{1-p_{00}}a_0(d);\\
P(F=d+j) &= \frac{1-(p_{00})^{c-1}}{1-p_{00}}\left[a_1(j)a_0(d) + p_0\frac{1-(p_{00})^{c-1}}{1-p_{00}}a_0(j)a_0(d)\right.\\
&\quad + \left.\sum_{k=1,\ k\neq j}^{d-1} P(F=d+j-k)a_0(k)\right] \quad \text{for } j = 1,2,\ldots,d-1.
\end{aligned}$$

The remaining probabilities can be calculated recursively

$$P(F=y) = \frac{1-(p_{00})^{c-1}}{1-p_{00}}\sum_{j=1}^{d-1} a_0(j)P(F=y-j) \quad \text{for} \quad y \geq 2d.$$

Example 2.2. Let us consider a random walk model over the non-negative integers determined by the following transition matrix

$$\begin{pmatrix}
p_{00} & p_{01} & 0 & 0 & 0 & \cdots\\
p_{10} & p_{11} & p_{12} & 0 & 0 & \cdots\\
0 & p_{21} & p_{22} & p_{23} & 0 & \cdots\\
0 & 0 & p_{32} & p_{33} & p_{34} & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{pmatrix}$$

and $p_0 + p_1 + \cdots = 1$.

If we substitute $d = 3$, we have the following expressions

$$\begin{aligned}
a_0(1) &= p_{01}p_{10};\\
a_0(2) &= p_{01}p_{11}p_{10};\\
a_0(3) &= p_{01}[p_{11}(p_{11}+p_{12}) + p_{12}];\\
a_1(1) &= p_1p_{10};\\
a_1(2) &= p_1p_{11}p_{10};\\
a_1(3) &= 1 + p_1[p_{11}(p_{11}+p_{12}) + p_{12} - 1] + p_2p_{21}(p_{11}+p_{12}-1) - p_0
\end{aligned}$$

and one can calculate the probabilities $P(F=y),\ y \geq 0$.

Remark 2.5. Letting $c \to \infty$ in the event $\mathbf{ST}(c,d)$ yields the event $\mathbf{ST}_2(d)$. From (3) we obtain the joint PGF $g(\infty,d)$ until a run of d consecutive failures in a multi-state Markov chain is observed, given by

$$g(\infty,d)=\sum_{g_1=1}^{\infty}\left[p_{g_1}t_{g_1}+\frac{p_0t_0+\sum_{l=1}^{\infty}p_lt_lF_{l0}^{(d-1)}}{1-p_{00}t_{00}\sum_{l=1}^{\infty}p_{0l}t_{0l}F_{l0}^{(d-1)}}p_{0g_1}t_{0g_1}\right]F_{g_1}^{(d)}, \qquad (11)$$

where $F_{g_1}^{(d)}$ and $F_{g_10}^{(d-1)}$ are the expressions (6) and (7), respectively.

As in the previous case, (11) is true only for irreducible Markov chains with persistent set of states.

3. LATER WAITING TIME PROBLEMS

The latter case in this section arises when stopping the experiment upon completion of both a run of c successes and a run of d failures in a multi state Markov chain, i.e. we deal with the event $\mathbf{LT}(c,d)$ and associated random variables

$$S^1_{g_1g_2} = \left\{\begin{array}{c}\text{number of transitions of type } g_1 \longrightarrow g_2 \text{ in the}\\ \text{Markov chain until the event } \mathbf{LT}(c,d) \text{ occurs}\end{array}\right\},$$

for $g_1, g_2 = 0,1,\ldots$ under condition that $P(\mathbf{LT}(c,d))=1$.

The joint PGF $g_{\mathbf{LT}}(c,d)$ related with the random vector

$$\mathbf{S}^1=(S_0,S_1,\ldots,S^1_{00},S^1_{01},\ldots)$$

can be obtained from the results derived in Section 2 because of duality noticed by Balasubramanian et *al.* (1993) and based on the relations (1) and (2).

If we denote by $g_{\mathbf{ST}}(c,d)$ the joint PGF given by (3), then the joint PGF $g_{\mathbf{LT}}(c,d)$ is determined by

$$g_{\mathbf{LT}}(c,d)=g(c,\infty)+g(\infty,d)-g_{\mathbf{ST}}(c,d), \qquad (12)$$

where $g(c,\infty)$ and $g(\infty,d)$ are expressions (10) and (11). Similar relationships hold for any marginal PGF's.

Remark 3.1. From (12) one can obtain the corresponding joint PGF, given by Aki and Hirano (1993) in their Theorem 2 where it is understood that X_0 is not counted in measuring the length of the run in two-state Markov chain.

Remark 3.2. The condition $P(\mathbf{LT}(c,d)) = 1$ implies that the joint PGF determined by (12) have to be modified as follows

$$g_{\mathbf{LT}}(c,d,c_1,d_1) = g(c,d_1) + g(c_1,d) - g_{\mathbf{ST}}(c,d),$$

where $c \leq c_1 < \infty$ and $d \leq d_1 < \infty$ with c_1 and d_1 being the maximal possible length of runs of successes and failures within the set $\mathbf{P}_0$, defined by Remark 2.1.

4. CONCLUSIONS

The sooner and later waiting time problems discussed by Balasubramanian et *al.* (1993) and Aki and Hirano (1993) have been generalized to the multi-state Markov chain. The obtained results can be applied for evaluating the system reliability (start-up demonstration tests), molecular biology (modeling the structure of DNA), for description of some random walk models, etc.

Presented results can also be extended for non-overlapping way of counting and for more complex sooner and latter waiting time problems considered by Uchida and Aki (1995) and Balakrishnan et *al.* (1997).

ACKNOWLEDGMENTS

This work is supported, in part, by the National Science Found of Bulgaria under Grant 705/97 and by FAPESP Grant 98/05810-9, Brazil.

BIBLIOGRAPHY

Aki, S., Balakrishnan, N. and Mohanty, S.G. (1996). Sooner and later waiting time problems for success and failure runs in higher order Markov dependent trials, *Ann. Inst. Stat. Math.*, **48**, 773-787.

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

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

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

Ebneshahrashoob, M. and Sobel, M. (1995). Dirichlet analysis for inverse multinomial with both quota and quota-free cells, *Computational Statistics & Data Analysis*, **19**, 293-307.

Fu, J.C. and Kuotras, M.V. (1996). Distribution theory of runs: A Markov chain approach, *J. Am. Statist. Assoc.*, **89**, 1050-1058.

Kolev, N. and Minkova, L. (1997). Discrete distributions related to success runs of length K in a multi-state Markov chain, *Communications in Statistics: Theory and Methods*, **26**, 1031-1049.

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

ÚLTIMOS RELATÓRIOS TÉCNICOS PUBLICADOS

9801 - GUIOL, H. Some properties of K-step exclusion processes. 1998. 21p. (RT-MAE-9801)

9802 - ARTES, R., JØRSENSEN, B. Longitudinal data estimating equations for dispersion models. 1998. 18p. (RT-MAE-9802)

9803 - BORGES, W.S.; HO, L.L. On capability index for non-nornal processes. 1998. 13p. (RT-MAE-9803)

9804 - BRANCO, M.D. , BOLFARINE, H., IGLESIAS, P., ARELLANO-VALLE, R. Bayesian analysis of the calibration problem under elliptical distributions. 1998. 15p. (RT-MAE-9804)

9805 - IGLESIAS, P., MATÚS, F., PEREIRA, C.A.B., TANAKA, N.I. On finite sequences conditionally uniform given minima and maxima. 1998. 13p. (RT-MAE-9805)

9806 - IGLESIAS, P., PEREIRA, C.A.B., TANAKA, N.I. Characterizations of multivariate spherical distribution in l_∞ norm. 1998. 19p. (RT-MAE-9806)

9807 - GIAMPAOLI, V., SINGER, J.M. A note on the comparison of two normal populations with restricted means. 1998. 12p. (RT-MAE-9807)

9808 - BUENO, V.C. A note on conditional reliability importance of components. 1998. 07p. (RT-MAE-9808)

9809 - GIMENEZ, P., BOLFARINE, H. Consistent estimation in comparative calibration models. 1998. 19p. (RT-MAE-9809)

9810 - ARTES, R., PAULA, G.A., RANVAUD, R. Estudo da direção tomada por pombos através de equações de estimação para dados circulares longitudinais. 1998. 16p. (RT-MAE-9810)

9811 - LIMA, A.C.P., SEM, P.K. Time-Dependent Coefficients in a Multi-Event Model for Survival Analysis. 1998. 21p. (RT-MAE-9811)

9812 - KOLEV, N., MINKOVA, L. Over-and underdispersed models for ruin probabilities. 1998. 34p. (RT-MAE-9812)

9813 - KOLEV, N. Negative Binomial Cross-Classifications. 1998. 16p. (RT-MAE-9813)

9814- **BAKEVA, V., GEORGIEVA, M., KOLEV, N.** Two Characterizations of the Geometric Distribution Related to an Unreliable Geo/G_D/1 System. 1998. 11p. (RT-MAE-9814)

9815 - **KOLEV, N., MINKOVA, L.** Some basic Inflated-Parameter Discrete Distributions. 1998. 19p. (RT-MAE-9815)

9816 - **ARTES, R., PAULA, G.A., RANVAUD, R.** Strategy for the analysis of circular longitudinal data based on generalized estimating equations. 1998. 15p. (RT-MAE-9816)

9817 - **PAULA, G.A., ARTES, R.** A Score-Type Test to Assess Overdispersion in Linear Logistic Models. 1998. 17p. (RT-MAE-9817)

9818 - **LOSCHI, R.H., IGLESIAS, P.L., WECHSLER, S.** Unpredictability and Probability Updating. 1998. 11p. (RT-MAE-9818)

9819 - **KOLEV, N., MINKOVA, L.** New Over-/Under- Dispersed Class of Inflated-Parameter Discrete Probability Distributions. 1998. 10p. (RT-MAE-9819)

9820 - **KOLEV, N., MINKOVA, L.** Inflated-Parameter Family of Generalized Power Series Distributions. 1998. 11p. (RT-MAE-9820)

9821 - **AUBIN, E.C.Q., STANGENHAUS, G.** A Robust Estimation of Growth Curves. 1998. 7p.(RT-MAE-9821)

9822 - **BORGES, W., HO, L. L.,** A Fraction Non-Conforming Based Capability Index. 1998. 13p. (RT-MAE-9822)

9823 - **GALEA-ROJAS, M., BOLFARINE, H.** Hypothesis Testing in Functional Comparative Calibration Models. 1998. 13p. (RT-MAE-9823)

9824 - **FONTES, L. R. G., ISOPI, M., NEWMAN, C. M.** Chaotic Time Dependence in a Disordered Spin System. 1998. 23p. (RT-MAE-9824)

9825 - **FERRARI, P. A., FONTES, L. R. G., VARES, M. E.** The Asymmetric Simple Exclusion Model With Multiple Shocks. 1998. 24p. (RT-MAE-9825)

9826 - **CABRAL, C. R. B., PAULA, G. A., BOLFARINE, H.** One-Sided Tests for the Analysis of Pretest-Posttest Data. 1998. 12p. (RT-MAE-9826)

9827 - **GASCO, L., BOLFARINE, H., ARELLANO-VALLE, R. B.** Asymptotic properties of the sample mean vector and sample covariance matrix with application in elliptical multiplicative measurement error models. 1998. 17p. (RT-MAE-9827)

9828 - **FERRARI, S. L. P., CRIBARI-NETO, F.** On the robustness of analytical and bootstrap corrections to score tests in regression models. 1998. 11p. (RT-MAE-9828)

9829 - **PERES, C. A., ANDRE, C. D. S., NARULA, S. C.** Whyn the observations within a treatment of a random-effects model are correlated? 1998. 8p. (RT-MAE-9829)

9830 - **FONTES, L.R., MATHIEU, P., PICCO, P.** On the averaged dynamics of the random field curie-weiss model. 1998. 27p. (RT-MAE-9830)

9831 - **ANDRÉ, C.D.S., ELIAN, S.N., TAVARES, R.** Coeficientes de determinação na regressão de mínima soma dos erros absolutos e sua utilização em problemas de seleção de variáveis. 1998. 14p. (RT-MAE-9831)

9832 - **SOUZA, F.A.M., PAULA, G.A.** Deviance residuals for an angular response. 1998. 13p. (RT-MAE-9832)

9833 - **SOLER, J.M.P., SINGER, J.M.** Optimal covariance adjustment in growth curve models. 1998. 13p. (RT-MAE-9833)

9834 - **BOTTER, D.A., PERES, A.A.Q., ZANDONADE, E.** A longitudinal study of mortality and air pollution for São Paulo, Brazil. 1998. 16p. (RT-MAE-9834)

9835 - **CRIBARI-NETO, F., FERRARI, S.L.P.** Monotonic improved critical values for econometric chi-squared asymptotic criteria. 1998. 12p. (RT-MAE-9835)

9836 - **CORDEIRO, G.M., FERRARI, S.L.P., URIBE-OPAZO, M.A., VASCONCELLOS, K.L.P.** Corrected Maximum Likelihood estimation in a class of symmetric nonlinear regression models[1]. 1998. 14p. (RT-MAE-9836)

9837 - **BELITSKY, V., FERRARI, P.A.** Invariant measures and convergence for cellular automaton 184 and related processes. 1998. 19p. (RT-MAE-9837)

9838 - **CASSANDRO, M., COLLET, P., GALVES, A., GALVES, CHARLOTTE.** A Statistical-Physics Aprroach to Language Acquisition and Language Change. 1998. 15p. (RT-MAE-9838)

9839 - **ANDRADE, D.F., VALLE, R.C.** Uma introdução à teoria da resposta ao item : conceitos e aplicações. 1998. 15p. (RT-MAE-9839)

9840 - **BUENO, V.C.** Component importance in a random environment. 1998. 6p. (RT-MAE-9840)

9841 - **OGLIARI, P.J., ANDRADE, D.F.** Analysing longitudinal data via nonlinear models in randomized block designs. 1998. 18p. (RT-MAE-9841)

9842 - **CANCHO, G.V., BOLFARINE, H., ACHCAR, J.A.** Use of the exponentiated-weibull distribution for bathtub-shaped failure rates. 1998. 25p. (RT-MAE-9842)

9901 - **MONTOYA-DELGADO, L.E., IRONY, T.Z., WHITTLE, M.R., PEREIRA, C.A.B.** Unconditional Exact Test of Hardy-Weinberg Law. 1999. 19p.(RT-MAE-9901)

9902 - **MINKOVA, L.** Modelling of financial markets in high level of inflation. 1999. 8p. (RT-MAE-9902)

9903 - **GUERBEROFF, G.R.** Remark on the critical temperature of the hopfield model. 1999. 10p. (RT-MAE-9903)

9904 - **BAKEVA, V., KOLEV, N.** The optimal blocking time of the unreliable Geo/G_D/1 Queuing System. 1999. 12p. (RT-MAE-9904)

The complete list of "Relatórios do Departamento de Estatística", IME-USP, will be sent upon request.

Departamento de Estatística
IME-USP
Caixa Postal 66.281
05315-970 - São Paulo, Brasil