On the frequency count for a random walk with absorbing boundaries: a carcinogenesis example. I

1994 J. Phys. A: Math. Gen. 27 7035

(http://iopscience.iop.org/0305-4470/27/21/018)

J. Phys. A: Math. Gen. 27 (1994) 7035–7046. Printed in the UK

On the frequency count for a random walk with absorbing boundaries: a carcinogenesis example. I

M A El-Shehawey

Department of Mathematics, Damietta Faculty of Science, New Damietta, Egypt

Received 23 March 1994, in final form 2 August 1994

Abstract. A non-homogeneous random walk on non-negative integers with transition probabilities $p_{0i} = \delta_{0i}$, $p_{Ni} = \delta_{Ni}$, $P_{i,i+1} = \lambda_i$, $p_{i,i-1} = \mu_i$, and $p_{i,i} = \rho_i$, $\lambda_i + \mu_i + \rho_i = 1$, is studied. In particular, when the transition probabilities are independent of position, a general expression for the joint probability generating function (JPGF) of the frequency count of the stages $1, 2, \ldots N-1$ is derived. The appropriate marginal forms of this JPGF yield the PGF of the frequency count at any pair of stages, and at any particular single stage. Some moment formulae associated with the frequency count are derived. A random walk conditional on absorption at a specified boundary is also considered. The random walk model proposed is eminently suitable for the example of carcinogenesis.

1. Introduction

Random walks with absorbing boundaries provide a natural model for a wide variety of phenomena that arise in medicine and biology. In this paper a random walk model of the phenomenon of carcinogenesis (see Bell (1976), and Beyer and Waterman (1979) and references cited there) is considered.

A tumour is an abnormal mass of tissue which is not inflammatory. A cancer tumour is usually thought of as arising from one wayward cell that has lost the ability to control itself. A cancer tumour inducing agent is called a carcinogen. In the study of carcinogenesis, a hit refers to the interaction between the carcinogen and the normal cell which results in the mutation of that normal cell to a cancer cell. The transition of a normal cell to a malignant cell need not occur in one hit or one stage. The number of stages is the number of mutations required to produce a cancer cell. A mutation is said to occur in a given stage if, during that stage, the mutated cell is subject to reproduction, death, further mutation to the next stage, etc. The natural model for this problem is a birth and death process with linear growth. This model has been extensively studied by many authors, perhaps more for its mathematical manageability than its genetic relevance, among which we may mention Bartlett (1960), Bharucha-Reid (1960), Gani and Jerwood (1971), Iosifescu and Tauta (1973), Bell (1976), Beyer and Waterman (1979), Adomian (1980), Iosifescu (1980), Karlin and Taylor (1984), Sumita (1984), Sumita and Masuda (1985), and Asmussen (1987).

In a multi-stage model one postulates several successive mutations, each producing a clone of mutant cells.

The assumptions made in the random walk model of carcinogenesis are:

(1) Let $\{X_n; n = 0, 1, \ldots\}$ denote the random walk corresponding to the mutation process, and $\{0, 1, \ldots, N\}$ denote the number of stages.

(2) The walk starts at stage $i \in I_{N-1} = \{1, 2, \ldots, N-1\}$.

0305-4470/94/217035+12$19.50

(3) A step forward implies further mutation to the next stage and a backward step implies a move towards recovery.

(4) The stage 0 represents the stage of complete recovery and stage N denotes the completion of the mutation process resulting in malignant cells.

(5) The random walk is governed by the one-step transition probability matrix $\mathbf{M} = (p_{ij})$, where

$$p_{ij} = \begin{cases} \lambda_i & j = i+1 \\ \mu_i & j = i-1 \quad i \in I_{N-1} \\ \rho_i & j = i \end{cases}$$

$\lambda_i + \mu_i + \rho_i = 1$, and $p_{0i} = \delta_{0i}$, $p_{Ni} = \delta_{Ni}$ (see figure 1).

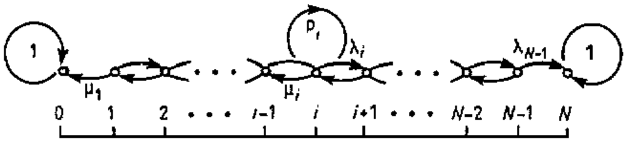

Figure 1. The state diagram of a non-homogeneous random walk with absorbing boundaries.

(6) $T_{i1}, T_{i2}, \ldots, T_{iN-1}$ are random variables denoting the frequency count (total number of occurrences) of the stages $1, 2, \ldots, N-1$, respectively, before entering one or the other boundary stage, given the initial stage $X_0 = i \in I_{N-1}$.

The purpose of this paper is to obtain a general formula for the JPGF of the frequency count for the non-homogeneous random walk. The appropriate marginal forms yield the PGF of the frequency count at any pair of stages, and at any particular single stage. When the spatial homogeneity is present explicit expressions for the corresponding JPGF are given. The covariance and the correlation coefficient of the frequency count at any pair of stages are calculated. Expressions are deduced for the distribution of a backward stage or a forward stage conditioned on hitting one of the boundaries before hitting the other. The probabilities conditioned on absorption at the origin of a homogeneous random walk are also given.

2. The JPGF of the frequency count for the non-homogeneous random walk

Let T_{ij} denote the random variable defined as the number of visits to stage j before eventual absorption at one of the boundaries (in other words the frequency count of j), given the starting stage i.

We introduce the following JPGF of the random variables $T_{i1}, T_{i2}, \ldots, T_{iN-1}$:

$$\begin{aligned} G_i(Z) = G_i(z_1, z_2, \ldots, z_{N-1}) &= E[z_1^{T_{i1}} z_2^{T_{i2}} \ldots z_{N-1}^{T_{iN-1}}] \\ &= \sum \mathrm{pr}\,(T_{i1} = n_1, T_{i2} = n_2, \ldots, T_{iN-1} = n_{N-1}) \prod_{j=1}^{N-1} z_j^{n_j} \\ & \quad |z_j| \leqslant 1, \; j \in I_{N-1} \end{aligned} \tag{2.1}$$

in which the summation extends over all $n_1, n_2, \ldots, n_{N-1}$ such that $\sum_{j\in I_{N-1}} n_j = t+1$, where t is interpreted as the number of transitions to either of the boundaries.

The master equation for the probability $\mathrm{pr}(T_{ik} = n_k,\ k \in I_{N-1})$ can be derived easily. The variables $\{n_k\}$ can be transformed to $\{z_k\}$ by generating function techniques. The resulting equations for the transform $G_i(Z)$ are given by the recursion

$$G_i(Z) = \frac{z_i}{1-\rho_i z_i}[\mu_i G_{i-1}(Z) + \lambda_i G_{i+1}(Z)] \qquad i \in I_{N-1} \tag{2.2}$$

subject to the boundary conditions

$$G_0(Z) = G_N(Z) = 1. \tag{2.3}$$

The above can be solved systematically, as described in theorem 2.1.

Theorem 2.1.

$$G_i(Z) = \frac{1}{1-F_{N-1}(Z)}\left[(1-F_{i-1}(Z))\prod_{j=1}^{N-i}\left(\frac{\lambda_{N-j}z_{N-j}}{1-\rho_{N-j}z_{N-j}}\right)\right.$$
$$\left. + (1-B_{N-i-1}(Z))\prod_{j=1}^{i}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right] \qquad i \in I_{N-1} \tag{2.4}$$

where $F_m(Z)$ and $B_m(Z)$ satisfy the recursion

$$F_m(Z) = F_{m-1}(Z) + \lambda_{m-1}\mu_m(1-F_{m-2}(Z))\prod_{j=m-1}^{m}\left(\frac{z_j}{1-\rho_j z_j}\right)$$
$$m = 2, 3, \ldots, N-1 \tag{2.5}$$

and

$$B_m(Z) = B_{m-1}(Z) + \lambda_{N-m}\mu_{N-m+1}(1-B_{m-2}(Z))\prod_{j=N-m}^{N-m+1}\left(\frac{z_j}{1-\rho_j z_j}\right)$$
$$m = 2, 3, \ldots, N-1. \tag{2.6}$$

Proof. Formula (2.2) can be reduced from second order to first order as follows: We start with

$$G_1(Z) = \frac{z_1}{1-\rho_1 z_1}[\mu_1 + \lambda_1 G_2(Z)]. \tag{2.7}$$

Inserting (2.7) into (2.2) immediately leads to

$$G_3(Z) = \frac{1-\rho_2 z_2}{\lambda_2 z_2(1-F_1(z_1))}\left[(1-F_2(Z))G_2(Z) - \prod_{j=1}^{2}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right] \tag{2.8}$$

where

$$F_2(Z) = F_2(z_1, z_2) = \lambda_1\mu_2\prod_{j=1}^{2}\frac{z_j}{1-\rho_j z_j} \qquad F_1(z_1) \equiv F_0 \equiv 0.$$

Inserting $G_2(Z)$ from (2.8) into (2.2), we obtain

$$G_4(Z) = \frac{1-\rho_3 z_3}{\lambda_3 z_3(1-F_2(Z))}\left[(1-F_3(Z))G_3(Z) - \prod_{j=1}^{3}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right] \tag{2.9}$$

where

$$F_3(Z) = F_3(z_1, z_2, z_3) = F_2(Z) + \lambda_2\mu_3(1-F_1(z_1))\prod_{j=2}^{3}\frac{z_j}{1-\rho_j z_j}.$$

Proceeding in the same fashion, we obtain

$$G_i(Z) = \frac{1-\rho_{i-1} z_{i-1}}{\lambda_{i-1} z_{i-1}(1-F_{i-2}(Z))}\left[(1-F_{i-1}(Z))G_{i-1}(Z) - \prod_{j=1}^{i-1}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right]$$
$$i \in I_{N-1} \tag{2.10}$$

where $F_m(Z)$ satisfies the recursion (2.5).

Evaluating $G_{N-2}(Z)$ from (2.10) and inserting the result into (2.2), we deduce that

$$G_N(Z) = \frac{1-\rho_{N-1} z_{N-1}}{\lambda_{N-1} z_{N-1}(1-F_{N-2}(Z))} \times \left[(1-F_{N-1}(Z))G_{N-1}(Z) - \prod_{j=1}^{N-1}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right]. \tag{2.11}$$

On account of the boundary condition given by (2.3), the expression (2.11) becomes

$$G_{N-1}(Z) = \frac{1}{1-F_{N-1}(Z)}\left[\frac{\lambda_{N-1} z_{N-1}}{1-\rho_{N-1} z_{N-1}}(1-F_{N-2}(Z)) + \prod_{j=1}^{N-1}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right]. \tag{2.12}$$

Reversing the stages, by setting $i = N-k$, $k \in I_{N-1}$ in (2.10), one finds that

$$G_{N-k-1}(z) = \frac{1}{1-F_{N-k-1}(Z)} \times \left[\frac{\lambda_{N-k-1} z_{N-k-1}}{1-\rho_{N-k-1} z_{N-k-1}}(1-F_{N-k-2}(Z))G_{N-k}(Z) + \prod_{j=1}^{N-k-1}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right]. \tag{2.13}$$

Inserting (2.12) into (2.13), we obtain

$$G_{N-2}(Z) = \frac{1}{1-F_{N-1}(Z)} \times \left[(1-F_{N-3}(Z))\prod_{j=1}^{2}\left(\frac{\lambda_{N-j} z_{N-j}}{1-\rho_{N-j} z_{N-j}}\right) + (1-B_1(z_{N-1}))\prod_{j=1}^{N-2}\left(\frac{\mu_j z_j}{1-\rho_j z_j}\right)\right] \tag{2.14}$$

where $B_1(z_{N-1}) \equiv B_0 \equiv 0$.

Substituting from (2.14) into (2.13), and using the fact that

$$1 - F_{N-1}(Z) + \lambda_{N-3}\mu_{N-2}(1 - F_{N-4}(Z)) \prod_{j=2}^{3} \left(\frac{z_{N-j}}{1 - \rho_{N-j} z_{N-j}} \right) = (1 - F_{N-3}(Z))(1 - B_2(Z)) \qquad (2.15)$$

where

$$B_2(Z) = B_2(z_{N-1}, z_{N-2}) = \lambda_{N-2}\mu_{N-1} \prod_{j=1}^{2} \frac{z_{N-j}}{1 - \rho_{N-j} z_{N-j}}$$

we obtain

$$G_{N-3}(Z) = \frac{1}{1 - F_{N-1}(Z)} \times \left[(1 - F_{N-4}(Z) \prod_{j=1}^{3} \left(\frac{\lambda_{N-j} z_{N-j}}{1 - \rho_{N-j} z_{N-j}} \right) + (1 - B_2(Z)) \prod_{j=1}^{N-3} \left(\frac{\mu_j z_j}{1 - \rho_j z_j} \right) \right]. \qquad (2.16)$$

Iterating further, we obtain (2.4), where $B_m(Z)$ satisfies the recursion (2.6). □

Many interesting probability generating functions can be derived from theorem 2.1 through an appropriate choice of the arguments z_j, $j \in I_{N-1}$. The next theorem follows immediately from (2.4) by setting all the arguments z_j equal to one, except z_x and z_y.

Theorem 2.2. The marginal PGF for two of the $N - 1$ random variables $T_{i1}, T_{i2}, \ldots, T_{iN-1}$ (say T_{ix}, T_{iy}) is given by

$$G_i(z_x, z_y) = \frac{1}{1 - f_{N-1}(z_x, z_y)} \left[r_1(1 - f_{i-1}(z_x, z_y)) \prod_{\substack{j=1 \\ j \neq x,y}}^{N-i} \left(\frac{\lambda_{N-j}}{1 - \rho_{N-j}} \right) + r_2(1 - b_{N-i-1}(z_x, z_y)) \prod_{\substack{j=1 \\ j \neq x,y}}^{i} \left(\frac{\mu_j}{1 - \rho_j} \right) \right] \qquad (2.17)$$

where

$$r_1 = \lambda_{N-x}\lambda_{N-y} \begin{cases} [(1 - \rho_{N-x})(1 - \rho_{N-y})]^{-1} & \text{if } x < i,\ y < i \\ [(1 - \rho_{N-x})(1 - \rho_{N-y} z_{N-y})]^{-1} z_{N-y} & \text{if } x < i,\ y \geqslant i \\ [(1 - \rho_{N-y})(1 - \rho_{N-x} z_{N-x})]^{-1} z_{N-x} & \text{if } x \geqslant i,\ y < i \\ [(1 - \rho_{N-x} z_{N-x})(1 - \rho_{N-y} z_{N-y})]^{-1} z_{N-x} z_{N-y} & \text{if } x \geqslant i,\ y \geqslant i \end{cases}$$

$$r_2 = \mu_x \mu_y \begin{cases} [(1 - \rho_x z_x)(1 - \rho_y z_y)]^{-1} z_x z_y & \text{if } x \leqslant i,\ y \leqslant i \\ [(1 - \rho_x)(1 - \rho_y z_y)]^{-1} z_y & \text{if } x > i,\ y \leqslant i \\ [(1 - \rho_y)(1 - \rho_x z_x)]^{-1} z_x & \text{if } x \leqslant i,\ y > i \\ [(1 - \rho_x)(1 - \rho_y)]^{-1} & \text{if } x > i,\ y > i \end{cases}$$

$f_m(z_x, z_y) = F_m(1, \ldots, 1, z_x, 1, \ldots, 1, z_y, 1, \ldots, 1)$, and $b_m(z_x, z_y) = B_m(1, \ldots, 1, z_x, 1, \ldots, 1, z_y, 1, \ldots, 1)$, $m = 0, 1, \ldots, N - 1$; z_x and z_y are the x and y components. □

The next corollary follows from theorem 2.2 by setting $z_y = 1$.

Corollary 2.1. The PGF of the total number of visits to stage x, T_{ix}, is given by

$$G_i(z_x) = \frac{1}{1-f_{N-1}(z_x)}\left[r_3(1-f_{i-1}(z_x))\prod_{\substack{j=1\\j\neq x}}^{N-i}\left(\frac{\lambda_{N-j}}{1-\rho_{N-j}}\right)\right.$$
$$\left. + \; r_4(1-b_{N-i-1}(z_x))\prod_{\substack{j=1\\j\neq x}}^{i}\left(\frac{\mu_j}{1-\rho_j}\right)\right] \tag{2.18}$$

where

$$r_3 = \lambda_{N-x}\begin{cases}(1-\rho_{N-x})^{-1} & \text{if } x < i\\ (1-\rho_{N-x}z_{N-x})^{-1}z_{N-x} & \text{if } x \geqslant i\end{cases}$$

$$r_4 = \mu_x\begin{cases}(1-\rho_x z_x)^{-1}z_x & \text{if } x \leqslant i\\ (1-\rho_x)^{-1} & \text{if } x > i\end{cases}$$

$f_m(z_x) = f_m(z_x, 1)$, and $b_m(z_x) = b_m(z_x, 1)$.

The recursion for $f_m(z_x)$ can be written as

$$f_m(z_x) = f_{m-1}(z_x) + \lambda_{m-1}\mu_m(1-f_{m-2}(z_x))K_m \tag{2.19}$$

where

$$K_m = \begin{cases}[(1-\rho_{m-1})(1-\rho_m z_m)]^{-1}z_m & m = x\\ [(1-\rho_{m-1}z_{m-1})(1-\rho_m)]^{-1}z_{m-1} & m = x+1\\ [(1-\rho_{m-1})(1-\rho_m)]^{-1} & m \neq x, x+1\end{cases}$$

and $b_m(z_x)$ satisfy the corresponding appropriate form of (2.6).

3. The JPGF of the frequency count for the homogeneous random walk

When spatial homogeneity is present, on setting $\lambda_i = \lambda$, $\mu_i = \mu$, $\rho_i = \rho$ for all $i \in I_{N-1}$, into (2.18) we obtain

$$G_i(z_x) = \frac{1}{1-h_{N-1}(z_x)}\left[a_1(1-h_{i-1}(z_x))\left(\frac{\lambda}{1-\rho}\right)^{N-i-1}\right.$$
$$\left. + \; a_2(1-g_{N-i-1}(z_x))\left(\frac{\mu}{1-\rho}\right)^{i-1}\right] \tag{3.1}$$

where

$$a_1 = \lambda\begin{cases}(1-\rho)^{-1} & x < i\\ (1-\rho z_x)^{-1}z_x & x \geqslant i\end{cases} \qquad a_2 = \mu\begin{cases}(1-\rho z_x)^{-1}z_x & x \leqslant i\\ (1-\rho)^{-1} & x > i\end{cases}$$

$h_m(z_x)$ satisfies the following recursion

$$h_m(z_x) = h_{m-1}(z_x) + \lambda\mu[1 - h_{m-2}(z_x)]A_m$$
$$A_m = \begin{cases} [(1-\rho)(1-\rho z_x)]^{-1} z_x & m = x, x+1 \\ (1-\rho)^{-2} & m \neq x, x+1. \end{cases} \tag{3.2}$$

It can be readily seen that the solution of (3.2) is given by

$$h_m(z_x) = \begin{cases} H_m & m < x \\ 1 - (1 - H_m)\left(\dfrac{1-\rho}{1-\rho z_x}\right) z_x & \\ \quad - (1 - H_{x-1})(1 - H_{m-x})\left(\dfrac{1 - z_x}{1 - \rho z_x}\right) & m \geqslant x \end{cases} \tag{3.3}$$

where H_m (independent of z_x) is the solution of the second order recursion

$$H_m = H_{m-1} + \frac{\lambda\mu}{(1-\rho)^2}[1 - H_{m-2}]$$
$$H_0 = H_1 = 0. \tag{3.4}$$

Similarly,

$$g_m(z_x) = \begin{cases} H_m & m < N - x \\ 1 - (1 - H_m)\left(\dfrac{1-\rho}{1-\rho z_x}\right) z_x & \\ \quad - (1 - H_{N-x-1})(1 - H_{m+x-N})\left(\dfrac{1 - z_x}{1 - \rho z_x}\right) & m \geqslant N - x. \end{cases} \tag{3.5}$$

Equation (3.4) can be solved employing standard methods, and we obtain

$$H_m = 1 - \frac{1}{(\lambda - \mu)(1-\rho)^m} \begin{cases} \lambda^{m+1} - \mu^{m+1} & \lambda \neq \mu \\ (\lambda - \mu)(m+1)\mu^m & \lambda = \mu. \end{cases} \tag{3.6}$$

One then sees from (3.3), (3.5) and (3.6) that

$$1 - h_{i-1}(z_x) = \frac{1}{\gamma_1^2(1-\rho)^{i-1}(1-\rho z_x)} \times \begin{cases} i\mu^{i-1}\gamma_1^2(1-\rho z_x) & i < x, \lambda = \mu \\ \mu^{i-2}(1-\rho)\gamma_1^2[x(i-x) + (i\mu - x(i-x))z_x] & i \geqslant x, \lambda = \mu \\ \gamma_1\gamma_i(1-\rho z_x) & i < x, \lambda \neq \mu \\ (1-\rho)[\gamma_x\gamma_{i-x} + (\gamma_1\gamma_i - \gamma_x\gamma_{i-x})z_x] & i \geqslant x, \lambda \neq \mu \end{cases} \tag{3.7}$$

and

$$1 - g_{N-i-1}(z_x) = \frac{1}{\gamma_1^2(1-\rho)^{N-i-1}(1-\rho z_x)}$$
$$\times \begin{cases} (N-i)\mu^{N-i-1}\gamma_1^2(1-\rho z_x) & i > x-1,\ \lambda = \mu \\ \mu^{N-i-2}(1-\rho)\gamma_1^2 \times [(N-x)(x-i) \\ \quad + ((N-i)\mu - (N-x)(x-i))z_x] & i \leqslant x-1,\ \lambda = \mu \\ \gamma_1\gamma_{N-i}(1-\rho z_x) & i > x-1,\ \lambda \neq \mu \\ (1-\rho)[\gamma_{N-x}\gamma_{x-i} \\ \quad + (\gamma_1\gamma_{N-i} - \gamma_{N-x}\gamma_{x-i})z_x] & i \leqslant x-1,\ \lambda \neq \mu \end{cases} \tag{3.8}$$

where $\gamma_p = \lambda^p - \mu^p$.

Inserting (3.7) and (3.9) into (3.1), one establishes the next theorem.

Theorem 3.1. Let $\lambda_i = \lambda$, $\mu_i = \mu$ and $\rho_i = \rho$, $i \in I_{N-1}$, where $\lambda + \mu + \rho = 1$. Then

$$G_i(z_x) = \frac{\alpha_{ix} + \beta_{ix} z_x}{\alpha_x - \beta_x z_x} \tag{3.9}$$

where

(1) for $\lambda \neq \mu$

$$\alpha_{ix} = \begin{cases} \lambda^{N-i}\gamma_x\gamma_{i-x} & x \leqslant i \\ \mu^i\gamma_{x-i}\gamma_{N-x} & x \geqslant i \end{cases}$$

$\beta_{ix} + \alpha_{ix} = \alpha_x - \beta_x = \gamma_1\gamma_N$, $\alpha_x = \gamma_x\gamma_{N-x}$, and $\gamma_p = \lambda^p - \mu^p$;

(2) for $\lambda = \mu$, $2\lambda + \rho = 1$,

$$\alpha_{ix} = \begin{cases} x(i-x) & x \leqslant i \\ (x-i)(N-x) & x \geqslant i \end{cases}$$

$\beta_{ix} + \alpha_{ix} = \alpha_x - \beta_x = \lambda N$, and $\alpha_x = x(N-x)$. □

Following the line of the proof of theorem 3.1, one establishes the next theorem.

Theorem 3.2. Let $\lambda_i = \lambda$, $\mu_i = \mu$ and $\rho_i = \rho$, $i \in I_{N-1}$, where $\lambda + \mu + \rho = 1$. The JPGF of the two dependent random variables T_{ix} and T_{iy} ($x, y \in I_{N-1}$) is

$$G_i(z_x, z_y) = \frac{\alpha_{i0} + \alpha_{i1}z_x + \alpha_{i2}z_y + \alpha_{i3}z_x z_y}{\alpha_0 + \alpha_1 z_x + \alpha_2 z_y + \alpha_3 z_x z_y} \tag{3.10}$$

where

$$\alpha_{i0} = \gamma_{y-x} \begin{cases} 0 & x \leqslant i \leqslant y \\ \mu^i\gamma_{x-i}\gamma_{N-y} & i \leqslant x \leqslant y \\ \lambda^{N-i}\gamma_{i-y}\gamma_x & x \leqslant y \leqslant i \end{cases}$$

$$\alpha_{i1} + \alpha_{i0} = \gamma_1 \begin{cases} \mu^i\gamma_{y-i}\gamma_{N-y} & x \leqslant i \leqslant y \text{ or } i \leqslant x \leqslant y \\ \lambda^{N-i}\gamma_{i-y}\gamma_y & x \leqslant y \leqslant i \end{cases}$$

$$\alpha_{i2} + \alpha_{i0} = \gamma_1 \begin{cases} \mu^i\gamma_{x-i}\gamma_{N-y} & i \leqslant x \leqslant y \\ \lambda^{N-i}\gamma_{i-x}\gamma_x & x \leqslant i \leqslant y \text{ or } x \leqslant y \leqslant i \end{cases}$$

$\alpha_0 = \gamma_x\gamma_{y-x}\gamma_{N-y}$, $\alpha_1+\alpha_0 = \gamma_1\gamma_y\gamma_{N-y}$, $\alpha_2+\alpha_0 = \gamma_1\gamma_x\gamma_{N-x}$, $\sum_{j=0}^{3}\alpha_{ij} = \sum_{j=0}^{3}\alpha_j = \gamma_1^2\gamma_N$, and $\gamma_p = \lambda^p - \mu^p$.

The case $y \leqslant x$ may be obtained by interchanging x and y in the above formulae.

□

It should be noted that theorem 3.1 follows directly from theorem 3.2 by setting $z_y = 1$, $\alpha_{i2}+\alpha_{i0} = \gamma_1\alpha_{ix}$, $\alpha_{i1}+\alpha_{i3} = \gamma_1\beta_{ix}$, $\alpha_2+\alpha_0 = \gamma_1\alpha_x$ and $\alpha_1+\alpha_3 = -\gamma_1\beta_x$.

Explicit expressions for the probability mass funciton (PMF) $\Pr(T_{ix} = n_x)$ can be calculated by expanding the denominator of (3.9) as a geometric series in $(\beta_x z_x/\alpha_x)$. We find

$$\Pr(T_{ix} = n_x) = \begin{cases} \dfrac{\alpha_{ix}}{\alpha_x} & n_x = 0 \\ \left[\dfrac{\alpha_{ix}}{\alpha_x} + \dfrac{\beta_{ix}}{\beta_x}\right]\left(\dfrac{\beta_x}{\alpha_x}\right)^{n_x} & n_x > 0. \end{cases} \tag{3.11}$$

It may be observed from (3.11) that $\Pr(T_{ix} = n_x)$ is geometric for $i = x$ (since $\alpha_{ix} = 0$ in this case), and modified geometric for $i \neq x$; α_x does not vanish, since $x \in I_{N-1}$ (see Barnett (1964), and El-Shehawey and Trabya (1993)).

4. The moment formulae for the homogeneous random walk

We derive some moment formulae associated with the two dependent random variables T_{ix} and T_{iy}, $x, y \in I_{N-1}$. These formulae are rather straightforward.

Theorem 4.1. The covariance, $\mathrm{Cov}(T_{ix}, T_{iy})$, and the correlation coefficient, Θ_{xy}, of T_{ix} and T_{iy} are given by

$$\mathrm{Cov}(T_{ix}, T_{iy}) = \frac{(\alpha_{i0}-\alpha_0)(\alpha_x-\beta_x) - (\alpha_{ix}\alpha_{iy} - \alpha_x\alpha_y)}{(\alpha_x-\beta_x)^2} \tag{4.1}$$

and

$$\Theta_{xy} = \frac{(\alpha_{i0}-\alpha_0)(\alpha_x-\beta_x) - (\alpha_{ix}\alpha_{iy} - \alpha_x\alpha_y)}{\sqrt{(\alpha_{ix}-\alpha_x)(\alpha_{iy}-\alpha_y)(\beta_{ix}-\alpha_x)(\beta_{iy}-\alpha_y)}} \tag{4.2}$$

where

$$\alpha_y = \frac{\alpha_1+\alpha_0}{\gamma_1} \qquad \alpha_{iy} = \frac{\alpha_{i1}+\alpha_{i0}}{\gamma_1} \qquad \beta_y = -\frac{\alpha_2+\alpha_3}{\gamma_1} \qquad \beta_{iy} = \frac{\alpha_{i2}+\alpha_{i3}}{\gamma_1}.$$

Finally, we may complete the discussion by obtaining explicit expressions for the distribution of a backward step or forward step conditioned by hitting one of the boundaries before hitting the other.

Let τ_j denote the first passage time to stage j, $j = 0, N$, i.e.

$$\tau_j = \begin{cases} \infty & \text{if } X_n \neq j \text{ for all } n,\ n = 1, 2, \ldots \\ \min(n \geqslant 1, X_n = j) & \text{if } X_n = j \text{ for some } n = 1, 2, \ldots. \end{cases} \tag{4.3}$$

Then the probabilities, $p_i(j)$, $j = 0, N$, for hitting one of the boundaries before hitting the other, given the initial position $X_0 = i$, (well known results) are

$$p_i(0) = \Pr(\tau_0 < \tau_N | X_0 = i) = \mu E[T_{i1}] = \begin{cases} \dfrac{\mu^i \gamma_{N-i}}{\gamma_N} & \lambda \neq \mu \\ \dfrac{N-i}{N} & \lambda = \mu \end{cases}$$

$$p_i(N) = \Pr(\tau_N < \tau_0 | X_0 = i) = \lambda E[T_{iN-1}] = \begin{cases} \dfrac{\lambda^{N-i} \gamma_i}{\gamma_N} & \lambda \neq \mu \\ \dfrac{i}{N} & \lambda = \mu \end{cases} \tag{4.4}$$

(see, for example, Feller (1977), Percus (1985), El-Shehawey (1992), and El-Shehawey and Trabya (1993)).

The probabilities for a backward step and forward step conditioned by hitting one of the boundaries before hitting the other are given by

$$\Pr(\tau_j = 1 | \tau_0 < \tau_N, X_0 = i) = \frac{p_{ij} p_j(0)}{p_i(0)}$$

$$\Pr(\tau_j = 1 | \tau_N < \tau_0, X_0 = i) = \frac{p_{ij} p_j(N)}{p_i(N)}. \tag{4.5}$$

One then sees from (4.4) and (4.5) that

$$\Pr(\tau_j = 1 | \tau_0 < \tau_N, X_0 = i) = \begin{cases} \dfrac{\gamma_{N-i+1}}{\gamma_{N-i}} & \lambda \neq \mu,\ j = i-1 \\ \dfrac{\lambda\mu\gamma_{N-i-1}}{\gamma_{N-i}} & \lambda \neq \mu,\ j = i+1 \\ \dfrac{\frac{1}{2}(1-\rho)(N-i+1)}{N-i} & \lambda = \mu,\ j = i-1 \\ \dfrac{\frac{1}{2}(1-\rho)(N-i-1)}{N-i} & \lambda = \mu,\ j = i+1 \end{cases} \tag{4.6}$$

and

$$\Pr(\tau_j = 1 | \tau_N < \tau_0, X_0 = i) = \begin{cases} \dfrac{\lambda\mu\gamma_{i-1}}{\gamma_i} & \lambda \neq \mu,\ j = i-1 \\ \dfrac{\gamma_{i+1}}{\gamma_i} & \lambda \neq \mu,\ j = i+1 \\ \dfrac{\frac{1}{2}(1-\rho)(i-1)}{i} & \lambda = \mu,\ j = i-1 \\ \dfrac{\frac{1}{2}(1-\rho)(i+1)}{i} & \lambda = \mu,\ j = i+1\,. \end{cases} \tag{4.7}$$

We might observe that (4.6) and (4.7) are invariant under interchange of λ and μ; also (4.7) can be deduced from (4.6) by interchanging λ and μ, interchanging $i-1$ and $i+1$, and replacing i by $N-i$. This invariance implies that, conditional on absorption at N, the distribution of time to absorption from $i \in I_{N-1}$ is the same as the distribution of time to absorption from $N-i$ conditional on absorption at 0.

The analogous results for a conditioned random walk with a single absorbing state at 0 may be immediately obtained as the limiting form of (4.6) when $N \to \infty$. In this case we obtain

$$\lim_{N\to\infty} \Pr(\tau_j = 1 | \tau_0 < \tau_N, X_0 = i) = \begin{cases} \max(\lambda, \mu) & \lambda \neq \mu,\ j = i-1 \\ \min(\lambda, \mu) & \lambda \neq \mu,\ j = i+1 \\ \frac{1}{2}(1-\rho) & \lambda = \mu,\ j = i-1 \\ \frac{1}{2}(\rho - 1) & \lambda = \mu,\ j = i+1. \end{cases} \tag{4.8}$$

5. Concluding remarks

The random walk with absorbing boundaries considered in this paper is perhaps deceptively simple. Whilst the general expression (2.4) for $G_i(Z)$ appears rather complicated in form, it is, nevertheless, consistent with intuitive ideas about the form of the JPGF of the total number of occurrences up to absorption (see Kemperman (1961), Barnett (1964), and El-Shehawey and Trabya (1993)).

The homogeneous random walk $\{X_n\}$ conditional on absorption at one of the boundaries yields a new non-homogeneous random walk with a state space $\{0, 1, \ldots, N\}$ and one-step transition probabilities:

$$q_{ij}(u) = \frac{p_{ij} p_j(u)}{p_i(u)} \qquad i, j \in I_{N-1} \quad u \in \{0, N\}. \tag{5.1}$$

In the case $\mu < \lambda$, formula (4.8) yields a new random walk of the same type in which forward and backward probabilities are interchanged.

Acknowledgment

The author would like to express his deepest appreciation to the anonymous referees for their constructive comments.

References

Adomian G 1980 *Applied Stochastic Processes* (New York: Academic)
Asmussen S 1987 *Applied Probability and Queues* (New York: Wiley)
Barnett V D 1964 *J. Austral. Math. Soc.* **4** 518–28
Bartlett M S 1960 *Population Models in Ecology and Epidemiology* (London: Methuen)
Bell G I 1976 *Science* **192** 569–72
Beyer W A and Waterman M A 1979 *Stud. Appl. Math.* **60** 83–90
Bharucha-Reid A T 1960 *Elements of the Theory of Markov Processes and Their Applications* (New York: McGraw-Hill)

El-Shehawey M A 1992 *Ann. Fac. Sci. Toulouse Ser. 6* **VI** 1–9

El-Shehawey M A and Trabya A M 1993 *Ann. Conf. Statistics, Computer Science, and Operational Research (Cairo University)* 28 91–103

Feller W 1977 *An Introduction to Probability Theory and Its Applications* vol 1 (New York: Wiley)

Gani J and Jerwood D 1971 *Biometrics* **27** 591–603

Iosifescu M and Tauta P 1973 *Stochastic Processes and Applications in Biology and Medicine* vol 2 (New York: Springer)

Iosifecu M 1980 *Finite Markov Processes and Their Applications* (New York: Wiley)

Karlin S and Taylor H M 1984 *A Second Course in stochastic Processes* (New York: Academic)

Kemperman J H B 1961 *The Passage Problem for a Stationary Markov Chain* (Chicago: University of Chicago Press)

Percus O E 1985 *Adv. Appl. Prob.* **17** 594–606

Sumita U 1984 *J. Appl. Prob.* **21** 10–21

Sumita U and Masuda Y 1985 *Stoch. Proc. Appl.* **20** 133–47