Exact solution for a two-stage stochastic evolution system

D. Arora, D. P. Bhatia, and M. A. Prasad
Division of Radiological Protection, Bhabha Atomic Research Centre, Trombay, Bombay 400 085, India
(Received 30 June 1992)

We obtain an analytical closed-form solution for a two-stage stochastic evolution system with time-independent transition rate parameters, which serves as a model for carcinogenesis. It is shown that the asymptotic probability of reaching the final stage never becomes unity in contrast to the deterministic stochastic model results available in literature. We also show that the first mutation event is more important than the second mutation.

PACS number(s): 02.50.−r, 05.40.+j

Stochastic evolution models find extensive use in many areas of physics, chemistry, and biology [1,2]. Here we consider a two-step model, incorporating birth and death processes. In this model a member of the system starting from stage S reaches the final stage M via an intermediate stage I (Fig. 1). Parameters α_n, β_n, and μ_n $(n=1,2)$ are the transition rates of the different pathways of the system. The number of members present in different stages determines the state of the system. Such a two-step process has been used as a model for carcinogenesis [3]. In this model S, I, and M represent the healthy, intermediate, and malignant stages, respectively, of a cell. D in Fig. 1 denotes a differentiated or a dead cell, which goes out of the system. Knudson [4] has shown that a few types of cancers have, in fact, just two stages of evolution. Henceforth we will use the language of "biological cells." This would, however, not restrict the use of this model in other similar areas of two-stage stochastic evolution systems. For detailed biological assumptions under which this model is applicable we refer the reader to Moolgavkar and Luebeck [5]. Such a model, where both stages evolve stochastically, has not been solved exactly so far. We present in this Brief Report a closed-form analytical solution of such a two-step system. The quantity of interest is the time-dependent probability of ending up with at least one malignant cell, given that we start with a single healthy cell at $t=0$.

Let $P_{i,j,k}(t)$ be the conditional probability of having exactly i healthy cells, j intermediate cells, and k malignant cells at any time t provided that, to start with, there is one normal cell and no intermediate and malignant cells at time $t=0$. The balance equation for the change of $P_{i,j,k}(t)$ with time is given by

$$\frac{dP_{ijk}}{dt}=(i-1)\alpha_1 P_{i-1,j,k}+(i+1)\beta_1 P_{i+1,j,k}+i\mu_1 p_{i,j-1,k}-i(\alpha_1+\beta_1+\mu_1)P_{i,j,k}$$
$$+(j-1)\alpha_2 P_{i,j-1,k}+(j+1)\beta_2 P_{i,j+1,k}+j\mu_2 P_{i,j,k-1}-j(\alpha_2+\beta_2+\mu_2)P_{i,j,k}\ . \tag{1}$$

By defining a probability generating function,

$$\Psi(X,Y,Z,t)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}P_{i,j,k}(t)X^iY^jZ^k\ ,$$

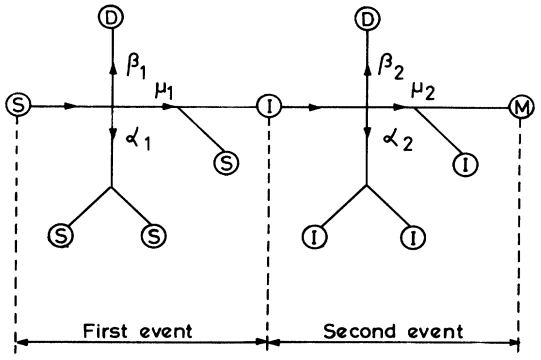

FIG. 1. A schematic representation of a two-stage stochastic evolution system.

multiplying it by Eq. (1) and carrying out the summation, we get

$$\Psi_t=\frac{d\Psi(X,Y,Z,t)}{dt}$$
$$=\{(\mu_1XY+\alpha_1X^2+\beta_1)-K_1X\}\frac{\partial\Psi}{\partial x}$$
$$+\{(\mu_2YZ+\alpha_2Y^2+\beta_2)-K_2Y\}\frac{\partial\Psi}{\partial Y} \tag{2}$$

where $K_n=\alpha_n+\beta_n+\mu_n$; $(n=1,2)$. This equation is to be solved with the initial condition

$$\Psi(X,Y,Z,0)=X\ .$$

Our quantity of interest, however, is $1-\Psi(1,1,0,t)$, which represents at any time t the probability of causation of malignancy, to be designated as $P(t)$.

CLOSED-FORM SOLUTION

Following the method of characteristics, the subsidiary equations from the partial differential equation (2) are

$$\frac{dZ}{dt}=0 , \tag{3}$$

$$\frac{dY}{dt}=K_2Y-(\mu_2YZ+\alpha_2Y^2+\beta_2) , \tag{4}$$

and

$$\frac{dX}{dt}=K_1X-(\mu_1XY+\alpha_1X^2+\beta_1) . \tag{5}$$

From Eq. (3), using $Z=\text{const}=0$ [since our interest is $\Psi(X,Y,0,t)$], Eq. (4) can easily be solved, and the constant trajectory solution in the (Y,t) plane is given by

$$A=[|Y-C_1|/|Y-C_2|]\exp(-pt) , \tag{6}$$

where A is the integration constant, $p=\alpha_2(C_2-C_1)$, and C_2 and C_1 are the two roots of $Y^2-(K_2/\alpha_2)Y+(\beta_2/\alpha_2)=0$.

By using the following transformation we define a new dependent variable U given by $X=(1/\alpha_1U)dU/dt$. From Eq. (6), using

$$Y=\frac{C_1+C_2A\exp(pt)}{1+A\exp(pt)} ,$$

Eq. (5) transforms into a second-order differential equation given by

$$\frac{d^2U}{dt^2}-\frac{dU}{dt}\frac{K_1-\mu_1(C_1+C_2Ae^{pt})}{1+Ae^{pt}}+\alpha_1\beta_1U=0 . \tag{7}$$

Now, a further transformation on t defined by $e^{pt}=z$ gives in place of Eq. (7) the new equation

$$p^2z^2(1+Az)\frac{d^2U}{dz^2}+pz(R+Sz)\frac{dU}{dz}+\alpha_1\beta_1(1+Az)U=0 , \tag{8}$$

where

$$R=p-K_1+\mu_1C_1$$

and

$$S=A(p-K_1+\mu_1C_2) .$$

Now using yet another transformation, $U=(z)^wy$, in Eq. (8), we get the differential equation

$$p^2z^2(1+Az)\frac{d^2y}{dz^2}+[2p^2wz(1+Az)+pz(R+Sz)]\frac{dy}{dz}$$
$$+[\{p^2Aw(w-1)+pSw+A\alpha_1\beta_1\}z+\{p^2w(w-1)+pwR+\alpha_1\beta_1\}]y=0 . \tag{9}$$

Now, in order to reduce this equation to a standard hypergeometric equation we impose the constraint

$$p^2w(w-1)+pwR+\alpha_1\beta_1=0 , \tag{10}$$

which implies two values of w, namely,

$$w_1=[(p-R)+\{(p-R)^2-4\alpha_1\beta_1\}^{1/2}]/2p ,$$

and

$$w_2=[(p-R)-\{(p-R)^2-4\alpha_1\beta_1\}^{1/2}]/2p . \tag{11}$$

With the constraint (10), Eq. (9) reduces to

$$p^2z(1+Az)\frac{d^2y}{dz^2}+[2p^2w(1+Az)+p(R+Sz)]\frac{dy}{dz}$$
$$+[p^2Aw(w-1)+pSw+A\alpha_1\beta_1]y=0 . \tag{12}$$

Now substituting $Az=-\theta$, Eq. (12) reduces to the standard hypergeometric equation

$$(\theta^2-\theta)\frac{d^2y}{d\theta^2}+\left[\frac{2wpA+S}{pA}\theta-\frac{2wp+R}{p}\right]\frac{dy}{d\theta}$$
$$+\left[w(w-1)+\frac{Sw}{pA}+\alpha_1\beta_1/p^2\right]y=0 . \tag{13}$$

The most general solution of Eq. (13) is given as

$$y=LF(a,b,c,\theta)+M\theta^{1-c}F(a-c+1,b-c+1,2-c,\theta)$$
$$=LF(1)+M\theta^{1-c}F(2) , \tag{14}$$

where F is the hypergeometric function. For brevity we will write $F(a,b,c,\theta)=F(1)$ and $F(a-c+1,b-c+1,2-c,\theta)=F(2)$. a, b, and c are given by the relations

$$a+b=2w+S/pA-1 , \tag{15}$$

$$ab=w(w-1)+Sw/pA+\alpha_1\beta_1/p^2 , \tag{16}$$

and

$$c=2w+R/p . \tag{17}$$

L and M in Eq. (14) are the integration constants to be determined by using the initial conditions. Now, restoring the transformations $\theta=-Az=-A\exp(pt)$ and $U=yz^w=y\exp(pwt)$ in Eq. (14), we get

$$U=\exp(pwt)M[(L/M)F(1)+\{-A\exp(pt)\}^{1-C}F(2)] , \tag{18}$$

or, in terms of the original variables X, Y, and t (where X was transformed as $X=\{dU/dt\}/\alpha_1U$, we get from Eq. (18)

$$\alpha_1X=\frac{dU}{dt}\Big/U$$
$$=pw+[NF(1)+\{-A\exp(pt)\}^{1-C}F(2)]^{-1}$$
$$\times\frac{d}{dt}[NF(1)+\{-A\exp(pt)\}^{1-C}F(2)] , \tag{19}$$

where $N=L/M$ is the new constant. From Eq. (19), N is given by

$$N=\frac{F'(2)\{-A\exp(pt)\}^{1-C}+F(2)\left[\frac{d}{dt}\{-A\exp(pt)\}^{1-C}\right]+[pw-\alpha_1 X][F(2)\{-A\exp(pt)\}^{1-C}]}{\alpha_1 XF(1)-F'(1)-F(1)pw} . \tag{20}$$

This is the required constant trajectory solution in the (X,t) plane corresponding to Eq. (5). Here the primes on hypergeometric functions $F(1)$ and $F(2)$ denote the differentiation with respect to time. Now in order to find the solution for the survival function $\Psi(X,Y,Z=0,t)$, we write, from Eq. (6),

$$A(t=0)=A_0=|Y-C_1|/|Y-C_2| , \tag{21}$$

and from Eq. (20),

$$N(t=0)=N_0=\frac{(-A_0)^{1-C}[F_0'(2)+F_0(2)\{p(w+1-c)-\alpha_1 X\}]}{(\alpha_1 X-pw)F_0(1)-F_0'(1)} . \tag{22}$$

(F_0, F_0', etc., are the values of the hypergeometric functions and their derivatives at $t=0$.)

Equation (22) gives

$$X=\frac{[N_0\{pwF_0(1)+F_0'(1)\}]+[F_0'(2)+F_0(2)(w+1-c)p](-A_0)^{1-c}}{N_0\alpha_1 F_0(1)+\alpha_1 F_0(2)(-A_0)^{1-C}} . \tag{23}$$

Now invoking initial condition $\Psi(X,Y,Z,0)=X$, we see that the actual functional form of $\Psi(X,Y,Z=0,0)$ is given by Eq. (23) itself. Now the survival function at any time t, i.e., $\Psi(X,Y,0,t)$, is found by restoring the time dependence in N_0 and A_0, i.e., by using N and A in place of N_0 and A_0 in Eq. (23). After a little algebra we get from Eq. (23)

$$\begin{aligned}\Psi(X,Y,0,t)=&(\exp\{p(1-c)t\}[pwF_0(1)+F_0'(1)][F'(2)+F(2)\{p(w+1-c)-\alpha_1 X\}]\\&-[F_0(2)(w+1-c)p+F_0'(2)][\{pw-\alpha_1 X\}F(1)+F'(1)])\\&\times(\exp\{p(1-c)t\}[\alpha_1 F_0(1)][F(2)\{(w+1-c)p-\alpha_1 X\}+F'(2)]\\&-[\alpha_1 F_0(2)][\{pw-\alpha_1 X\}F(1)+F'(1)])^{-1} .\end{aligned} \tag{24}$$

This is the exact closed-form solution of the two-stage stochastic-stochastic model. The quantity of interest, $P(t)$, can now be obtained from

$$P(t)=1-\Psi(1,1,0,t) . \tag{25}$$

Taking $t\to\infty$ as the limit in Eq. (25) (using proper limiting values of the hypergeometric functions and their derivatives) we find that $P(t)\sim 1-P\omega_2/\alpha_1$, i.e., the probability of reaching the final stage (malignancy in the biological case) never becomes 1, in contrast to the approximate results of Moolgavkar, Dewanji, and Venzon [6] or the deterministic stochastic model results of Quinn [7]. Their asymptotic value of $P(t)$ is always 1. Another important feature to be seen in the present work is that the first event (mutation) is relatively the more important of the two successive events in the sense that the asymptotic probability for small μ_1 depends upon α_1 and β_1 only [since for small μ_1 the quantity $P\omega_2/\alpha_1$ is β_1/α_1 and $P(t)\sim 1-\beta_1/\alpha_1$].

[1] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).
[2] T. J. Bailey, *The Elements of Stochastic Processes* (Wiley, New York, 1964).
[3] S. Moolgavkar and D. J. Venzon, Math. Biosci. **47**, 55 (1979).
[4] A. G. Knudson, Proc. Natl. Acad. Sci. U.S.A. **68**, 820 (1971).
[5] S. Moolgavkar and G. Luebeck, Risk Anal. **10**, 323 (1990).
[6] S. Moolgavkar, A. Dewanji, and D. J. Venzon, Risk Anal. **8**, 383 (1988).
[7] D. W. Quinn, Risk Anal. **9**, 407 (1990).