COLLOQUIUM MATHEMATICUM

VOL. XIII 1965 FASC. 2

A METHOD FOR DERIVATION OF PROBABILITIES IN A STOCHASTIC MODEL OF POPULATION GROWTH FOR CARCINOGENESIS

BY

W. KLONECKI (POZNAŃ)

1. Introduction. In the study of carcinogenesis (Neyman [3]) there arises the problem of deriving the probabilities describing a stochastic model of population growth. The probabilities cannot be deduced by the well-known procedure, due to Bartlett [1], which consists in replacing the system of differential equations by a partial differential equation, because the equation cannot be solved by standard methods. Since the problem is important, an effort has been made to find some other method.

There are two possible ways to handle cases where Bartlett's method of replacing the differential equation does not provide a solution. The first is to look for approximate solutions of the partial differential equation for the generating function. The second, which is used in this paper, is to look directly for exact formulae for the probabilities describing the stochastic model. The method used in this paper is an extension of Bartlett's method [1].

The second section of this paper considers two partial differential equations, which arise in the examples considered in the last section. The purpose is to deduce formulae for the coefficients of the expansions of the functions satisfying the differential equations. It has been shown that the partial infinite sums (see (7) and (51)) of the expansions of the functions satisfying the partial differential equations are completely determined by the first $n+2$ and $n+r+2$ coefficients. This fact simplifies the computation of the coefficients that we are seeking. General expressions are found for the partial sums of the expansions of functions satisfying the differential equations (theorem 1 and corollary 3) and relations between them are established (theorem 2 and corollary 4).

In the last section the results obtained in the second section are applied to solve problems concerned with one and two stage birth and death processes with constant rates. We find the probability that at time t there will be n live cells with m deaths and r mutations in the time interval from 0 to t. The main result is a procedure that can be used to evaluate approximately the joint distribution arising in Neyman's stochastic model of carcinogenesis.

I wish to acknowledge my indebtedness to Professor Jerzy Neyman for proposing this investigation to me and for his many valuable suggestions during its progress.

2. The partial differential equation. In this section the differential equation

$$\frac{\partial G}{\partial t} - [\lambda u^2 - \beta u + \gamma(t) v] \frac{\partial G}{\partial u} = 0 \tag{1}$$

coupled with the boundary condition $G(u, v, 0) = u$ will be considered. The coefficients λ and β are assumed to be constant. Our problem is to find the formulae for the coefficients $p_{m,n}(t)$ of the expansion

$$G(u, v, t) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p_{m,n}(t) u^m v^n \tag{2}$$

of the function $G(u, v, t)$ satisfying equation (1).

The partial differential equation (1) cannot be solved by standard methods, because they lead to a differential equation of the Riccati type.

If relation (2) is substituted into equation (1), then, by comparing appropriate terms, we obtain the system

$$p'_{m,n}(t) = \lambda(m-1) p_{m-1,n}(t) + \gamma(t)(m+1) p_{m+1,n-1}(t) - \beta m p_{m,n}(t) \tag{3}$$

of differential equations. The boundary condition $G(u, v, 0) = u$ reduces to $p_{1,0}(0) = 1$ and $p_{m,n}(0) = 0$ if $(m, n) \neq (1, 0)$. The functions $p_{m,n}(t)$ are defined to be zero for $m < 0$ or $n < 0$. This assumption is made throughout the paper. The differential equations (3) are of a recursive character and can be solved successively. This procedure can be however simplified. From theorem 1 below follows, as it will be shown, that if the first $n+2$ functions $p_{0,n}(t), p_{1,n}(t), \ldots, p_{n+1,n}(t)$ are known, then the functions $p_{n+2,n}(t), p_{n+3,n}(t), \ldots$ can be easily computed.

Introducing the notation

$$G_n(u, t) = \sum_{m=0}^{\infty} p_{m,n}(t) u^m \tag{4}$$

leads to the equation

$$\frac{\partial G_n}{\partial t} - (\lambda u^2 - \beta u) \frac{\partial G_n}{\partial u} = \gamma(t) \frac{\partial G_{n-1}}{\partial u} \tag{5}$$

and to the boundary conditions $G_0(u, 0) = u$, $G_1(u, 0) = 0$, $G_2(u, 0) = 0, \ldots$

THEOREM 1. *For every n, there exists a system of $n+2$ functions*

$$\psi_0^{(n)}(t), \psi_1^{(n)}(t), \ldots, \psi_{n+1}^{(n)}(t) \tag{6}$$

depending only on β and $\gamma(t)$, such that the partial sums (4) of the function $G(u, v, t)$, that satisfies (1), can be presented in the form

$$G_n(u, t) = \left(\frac{\lambda}{\beta}\right)^{n-1} \frac{\psi_0^{(n)}(t) + \frac{\lambda}{\beta}\psi_1^{(n)}(t)u + \ldots + \left(\frac{\lambda}{\beta}\right)^{n+1}\psi_{n+1}^{(n)}(t)u^{n+1}}{\left[1 - \frac{\lambda}{\beta}(1-\exp\{-\beta t\})u\right]^{n+1}}. \tag{7}$$

Proof. If $n = 0$, then (5) reduces to

$$\frac{\partial G_0}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial G_0}{\partial u} = 0, \tag{8}$$

with the boundary condition $G_0(u, 0) = u$. The solution of (8) can be found by the well-known standard method. According to this method first the differential equation

$$\frac{dt}{1} = \frac{du}{-(\lambda u^2 - \beta u)} \tag{9}$$

is to be solved. Using the solution of (9), which is equal to

$$u = \frac{\beta}{\lambda + k\beta e^{-\beta t}}, \tag{10}$$

where k is a constant, the function $G_0(u, t)$ satisfying (8) and the boundary condition $G_0(u, 0) = u$ can be obtained. Namely we have

$$G_0(u, t) = \frac{u\beta}{\beta\exp\{\beta t\} + \lambda u(1 - \exp\{\beta t\})}, \tag{11}$$

or

$$G_0(u, t) = \left(\frac{\lambda}{\beta}\right)^{-1} \frac{\psi_0^{(0)}(t) + \left(\frac{\lambda}{\beta}\right)\psi_1^{(0)}(t)u}{1 - \frac{\lambda}{\beta}(1 - \exp\{-\beta t\})u}, \tag{12}$$

where $\psi_0^{(0)}(t) = 0$ and $\psi_1^{(0)}(t) = e^{-\beta t}$. Hence relation (7) holds for $n = 0$.

If $n > 0$, we obtain by standard methods the recursive formulae

$$G_n(u, t) = \int_0^t \gamma(t_0)\left[\frac{\partial G_{n-1}(u, t_0)}{\partial u}\right]_{u\sim\beta(\lambda+k\beta e^{-\beta t_0})^{-1}} dt_0 \tag{13}$$

for the solution of (5). Here k has to be taken from relation (10). Eliminating k leads to

$$\beta(\lambda+k\beta e^{-\beta t_0})^{-1} = G_0(u, t-t_0). \tag{14}$$

It is also more convenient to write t_n instead of t and t_{n-1} instead of t_0. So we get the equation

$$G_n(u, t_n) = \int_0^{t_n} \gamma(t_{n-1})\left[\frac{\partial G_{n-1}(u, t_{n-1})}{\partial u}\right]_{u\sim G_0(u, t_n-t_{n-1})} dt_{n-1}. \tag{15}$$

Now we prove by induction the formula

$$G_n(u, t_n) = \int_0^{t_n} \gamma(t_{n-1}) \int_0^{t_{n-1}} \gamma(t_{n-2})\ldots\int_0^{t_1} \gamma(t_0)\, W_n(u, t_n, \ldots, t_0)\,dt_0\ldots dt_{n-1}, \tag{16}$$

where $W_n(u, t_n, \ldots, t_0)$ is a sum of expressions of the form

$$A_n\left[\frac{G_0(u, t_n)}{G_0(u, t_n-t_0)}\right]^{\varepsilon_0}\left[\frac{G_0(u, t_n)}{G_0(u, t_n-t_1)}\right]^{\varepsilon_1}\cdots\left[\frac{G_0(u, t_n)}{G_0(u, t_n-t_{n-1})}\right]^{\varepsilon_{n-1}}, \tag{17}$$

where ε_i equal $0, 1$ or 2, while

$$\sum_{i=0}^{n-1}\varepsilon_i = n+1.$$

Also A_n does not depend on u. The particular terms of W_n differ only in the A_n and the $(\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_{n-1})$.

If $n = 1$, then (15) reduces to

$$G_1(u, t_1) = \int_0^{t_1} \gamma(t_0)\left[\frac{\partial G_0(u, t_0)}{\partial u}\right]_{u\sim G_0(u, t_1-t_0)} dt_0. \tag{18}$$

Note that $G_0(u, t)$ given by (11) has the following properties:

$$G_0[G_0(u, t_1), t_2] = G_0(u, t_1+t_2), \tag{19}$$

$$\frac{\partial G_0(u, t)}{\partial u} = \left[\frac{G_0(u, t)}{u}\right]^2 e^{\beta t}. \tag{20}$$

Using (20) and (19) leads to

$$G_1(u, t_1) = \int_0^{t_1} \gamma(t_0)e^{\beta t_0}\left[\frac{G_0(u, t_1)}{G_0(u, t_1-t_0)}\right]^2 dt_0. \tag{21}$$

Hence $G_1(u, t_1)$ is of the form (16) with $A_1 = e^{\beta t_0}$ and $\varepsilon_0 = 2$.

Now it is to be shown that if formula (16) holds for n, then it holds also for $n+1$, i. e., we have

$$(22)\quad G_{n+1}(u,t_{n+1}) = \int_0^{t_{n+1}} \gamma(t_n)\int_0^{t_n}\gamma(t_{n-1})\ldots\int_0^{t_1}\gamma(t_0)W_{n+1}(u,t_{n+1},\ldots,t_0)\,dt_0\ldots dt_n,$$

where $W_{n+1}(u, t_{n+1}, \ldots, t_0)$ is a sum of expressions of an analogous form as (17), i. e., of the form

$$(23)\quad A_{n+1}\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_0)}\right]^{\eta_0}\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_1)}\right]^{\eta_1}\cdots\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_n)}\right]^{\eta_n},$$

where η_i equal $0, 1$ or 2 and

$$\sum_{i=0}^{n}\eta_i = n+2.$$

Using (15) and (16) leads to

$$(24)\qquad G_{n+1}(u,t_{n+1}) = \int_0^{t_{n+1}}\gamma(t_n)\left[\frac{\partial G_n(u,t_n)}{\partial u}\right]_{u\sim G_0(u,\,t_{n+1}-t_n)} dt_n =$$

$$\int_0^{t_{n+1}}\gamma(t_n)\left[\frac{\partial}{\partial u}\int_0^{t_n}\gamma(t_{n-1})\ldots\int_0^{t_1}\gamma(t_0)W_n(u,t_n,\ldots,t_0)\,dt_0\ldots dt_{n-1}\right]_{u\sim G_0(u,\,t_{n+1}-t_n)} dt_n$$

$$=\int_0^{t_{n+1}}\gamma(t_n)\int_0^{t_n}\gamma(t_{n-1})\ldots\int_0^{t_1}\gamma(t_0)\left[\frac{\partial W_n(u,t_n,\ldots,t_0)}{\partial u}\right]_{u\sim G(u,\,t_{n+1}-t_n)} dt_0\ldots dt_n.$$

The derivative of $W_n = W_n(u, t_n, \ldots, t_0)$ with respect to u is the sum of derivatives with respect to u of the particular terms (17). Therefore, in order to prove relation (22), it is sufficient to prove the following: by differentiating the particular terms (17) with respect to u and substituting $G_0(u, t_{n+1}-t_n)$ instead of u terms having the form of formula (23) are obtained. This follows from the relations

$$(25)\qquad \frac{\partial}{\partial u}\left[\frac{G_0(u,t_n)}{G_0(u,t_n-t_k)}\right]_{u\sim G_0(u,\,t_{n+1}-t_n)} = \frac{\lambda}{\beta}\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_n)}\right]^2(1-e^{-\beta t_k})e^{\beta t_n}$$

and

$$(26)\qquad \frac{\partial}{\partial u}\left[\frac{G_0(u,t_n)}{G_0(u,t_n-t_k)}\right]^2_{u\sim G_0(u,\,t_{n+1}-t_n)}$$

$$= 2\frac{\lambda}{\beta}\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_k)}\right]\left[\frac{G_0(u,t_{n+1})}{G_0(u,t_{n+1}-t_n)}\right]^2(1-e^{-\beta t_k})e^{\beta t_n}.$$

Both the formulae can be obtained by elementary computations using (20) and (19).

Note that

$$(27)\qquad \frac{G_0(u, t_n)}{G_0(u, t_n - t_k)} = \frac{1 - \frac{\lambda}{\beta}\big(1 - \exp\{-\beta(t_n - t_k)\}\big)u}{1 - \frac{\lambda}{\beta}\ (1 - \exp\{-\beta t_n\})u} e^{-\beta t_k}.$$

Hence (17) reduces to

$$(28)\qquad A_n \exp\Big\{-\beta \sum_{i=0}^{n-1} \varepsilon_i t_i\Big\} \times$$

$$\times \frac{\Big[1 - \frac{\lambda}{\beta}\big(1 - \exp\{-\beta(t_n - t_0)\}\big)u\Big]^{\varepsilon_0} \dots \Big[1 - \frac{\lambda}{\beta}\big(1 - \exp\{-\beta(t_n - t_{n-1})\}\big)u\Big]^{\varepsilon_{n-1}}}{\Big[1 - \frac{\lambda}{\beta}(1 - \exp\{-\beta t_n\})u\Big]^{n+1}},$$

where ε_i equal $0, 1$ or 2, while

$$\sum_{i=0}^{n-1} \varepsilon_i = n+1.$$

Whatever are the numbers $(\varepsilon_0, \varepsilon_1, \dots, \varepsilon_{n-1})$ expression (28) is a ratio of two polynomials of degree $n+1$ with respect to u with the denominator

$$\Big[1 - \frac{\lambda}{\beta}(1 - \exp\{-\beta t_n\})u\Big]^{n+1}.$$

Hence $W_n(u, t_n, \dots, t_0)$ is also a ratio of two polynomials of degree $n+1$ with the same denominator. In other words, formula (7) holds.

As it has been already mentioned, theorem 1 gives a way to evaluate the probabilities $p_{0,n}(t), p_{1,n}(t), \dots$ Using the equation

$$(29)\qquad \frac{1}{(1-u)^{n+1}} = \sum_{i=0}^{\infty} \binom{n+i}{i} u^i,$$

formula (7) can be presented in the form

$$(30)\qquad G_n(u, t) = \Big(\frac{\lambda}{\beta}\Big)^{n-1} \sum_{m=0}^{\infty} \Big(\frac{\lambda}{\beta}\Big)^m (1 - e^{-\beta t})^m \sum_{i=0}^{j} \binom{m+n-i}{n} \frac{\psi_i^{(n)}(t)}{(1 - e^{-\beta t})^i} u^m,$$

where $j = \min(m, n+1)$. Hence

$$(31)\qquad p_{m,n}(t) = \left(\frac{\lambda}{\beta}\right)^{m+n-1}(1-e^{-\beta t})^m \sum_{i=0}^{j}\binom{m+n-i}{n}\frac{\psi_i^{(n)}(t)}{(1-e^{-\beta t})^i},$$

where $j = \min(m, n+1)$. Also, if the first $n+2$ probabilities $p_{0,n}(t)$, $p_{1,n}(t), \ldots, p_{n+1,n}(t)$ are known, the functions $\psi_1^{(n)}(t), \ldots, \psi_{n+1}^{(n)}(t)$ can be evaluated from (31).

In the particular case when $\gamma(t)$ is constant, two corollaries can be obtained from theorem 1.

COROLLARY 1. *If $\gamma(t)$ is constant, the partial sums $G_n(u, t)$ of the expansion of the function $G(u, v, t)$ that satisfies* (1) *can be presented in the form*

$$(32)\qquad G_n(u, t)$$
$$= \left(\frac{\gamma}{\beta}\right)^n \left(\frac{\lambda}{\beta}\right)^{n-1} \frac{\varphi_0^{(n)}(\beta t) + \left(\frac{\lambda}{\beta}\right)\varphi_1^{(n)}(\beta t)u + \ldots + \left(\frac{\lambda}{\beta}\right)^{n+1}\varphi_{n+1}^{(n)}(\beta t)u^{n+1}}{\left[1-\frac{\lambda}{\beta}(1-\exp\{-\beta t\})u\right]^{n+1}},$$

where $\varphi_0^{(n)}(t), \ldots, \varphi_{n+1}^{(n)}(t)$ are functions of t only.

Proof. If $\gamma(t)$ is constant, then in view of the form of formulae (16) and (28), formula (7) can be written in the form (32).

COROLLARY 2. *If $\gamma(t)$ is constant, then the coefficients $p_{m,n}(t)$ of the expansion of the function $G(u, v, t)$ that satisfies* (1) *can be presented in the form*

$$(33)\qquad p_{m,n}(t) = \left(\frac{\gamma}{\beta}\right)^n \left(\frac{\lambda}{\beta}\right)^{m+n-1}(1-e^{-\beta t})^m \sum_{i=0}^{j}\binom{m+n-i}{n}\frac{\varphi_i^{(n)}(\beta t)}{(1-e^{-\beta t})^i},$$

where $j = \min(m, n+1)$.

Proof. Formula (33) is obtained from formula (32) in an analogous way as formula (31) from (7).

A possible way to evaluate the functions $\varphi_0^{(n)}(t), \ldots, \varphi_{n+1}^{(n)}(t)$ appearing in the expressions (32) and (33) is to use an integral representation of the form (16).

The other partial differential equation to consider is

$$(34)\qquad \frac{\partial G}{\partial t} - \big(\lambda u^2 - \beta u + \delta(t)v + \eta(t)w\big)\frac{\partial G}{\partial u} = 0,$$

with the associated boundary condition $G(u, v, w, 0) = u$. The coefficients λ and β are assumed to be constant.

As in the previous case, the formulae for the coefficients of the expansion

$$G(u, v, w, t) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{r=0}^{\infty} p_{m,n,r}(t)\, u^m v^n w^r \tag{35}$$

of the function $G(u, v, w, t)$ satisfying equation (34) are to be found.

In the same way as before, relation (34) can be split into a system of recurrent partial differential equations

$$\frac{\partial G_{n,r}}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial G_{n,r}}{\partial u} = \delta(t)\frac{\partial G_{n-1,r}}{\partial u} + \eta(t)\frac{\partial G_{n,r-1}}{\partial u}, \tag{36}$$

where

$$G_{n,r}(u, t) = \sum_{m=0}^{\infty} p_{m,n,r}(t)\, u^m, \tag{37}$$

with the boundary conditions $G_{0,0}(u, 0) = u$, $G_{n,r}(u, 0) = 0$ for $n^2+r^2 > 0$, and

$$G(u, v, w, t) = \sum_{n=0}^{\infty}\sum_{r=0}^{\infty} G_{n,r}(u, t)\, v^n w^r. \tag{38}$$

THEOREM 2. *For every n and r there is a system of $\binom{n+r}{r}$ sequences*

$$\gamma_1^{(s)}(t), \gamma_2^{(s)}(t), \ldots, \gamma_{n+r}^{(s)}(t), \tag{39}$$

where the $\gamma_i^{(s)}(t)$ equal $\delta(t)$ or $\eta(t)$, $1 \leqslant s \leqslant \binom{n+r}{r}$, such that the partial sum $G_{n,r}(u, t)$ of the expansion of the function $G(u, v, w, t)$ that satisfies (34) is equal to the sum of solutions $H_{n+r}^{(s)}(u, t)$ of the $\binom{n+r}{r}$ recurrent systems of the partial differential equations

$$\begin{gathered}
\frac{\partial H_0^{(s)}}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial H_0^{(s)}}{\partial u} = 0, \\
\frac{\partial H_1^{(s)}}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial H_1^{(s)}}{\partial u} = \gamma_1^{(s)}\frac{\partial H_0^{(s)}}{\partial u}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\frac{\partial H_{n+r}^{(s)}}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial H_{n+r}^{(s)}}{\partial u} = \gamma_{n+r}^{(s)}\frac{\partial H_{n+r-1}^{(s)}}{\partial u},
\end{gathered} \tag{40}$$

coupled with the boundary conditions

$$H_0^{(s)}(u, 0) = u, \quad H_1^{(s)}(u, 0) = H_2^{(s)}(u, 0) = \ldots = H_{n+r}^{(s)}(u, 0) = 0 \tag{41}$$

and corresponding to the particular sequences (39).

Proof. The theorem is true if one of the indices n or r is equal to zero, because then equation (36) is of the form of equation (5). Suppose now that the theorem holds for $(n-1, r)$ and $(n, r-1)$. By applying the standard method the solution of (36) is found to be

$$G_{n,r}(u,t)=\int_0^t \delta(t_0)\left[\frac{\partial G_{n-1,r}(u,t_0)}{\partial u}\right]_{u\sim\beta(\lambda+k\beta e^{-\beta t_0})^{-1}} dt_0+ \\ +\int_0^t \eta(t_0)\left[\frac{\partial G_{n,r-1}(u,t_0)}{\partial u}\right]_{u\sim\beta(\lambda+k\beta e^{-\beta t_0})^{-1}} dt_0, \tag{42}$$

where k has to be taken from relation (10). Using (14) leads to

$$G_{n,r}(u,t)=\int_0^t \delta(t_0)\left[\frac{\partial G_{n-1,r}(u,t_0)}{\partial u}\right]_{u\sim G_0(u,t-t_0)} dt_0+ \\ +\int_0^t \eta(t_0)\left[\frac{\partial G_{n,r-1}(u,t_0)}{\partial u}\right]_{u\sim G_0(u,t-t_0)} dt_0. \tag{43}$$

By assumption there is, for the function $G_{n-1,r}(u,t)$, a system of $\binom{n+r-1}{r}$ sequences

$$\gamma_1^{(s)}(t),\ \gamma_2^{(s)}(t),\ \ldots,\ \gamma_{n+r-1}^{(s)}(t), \tag{44}$$

where the $\gamma_i^{(s)}$ equal $\delta(t)$ or $\eta(t)$, $1\leqslant s\leqslant\binom{n+r}{r}$, such that we have

$$G_{n-1,r}(u,t)=\sum_{s=1}^{\binom{n+r-1}{r}} H_{n+r-1}^{(s)}(u,t), \tag{45}$$

where $H_{n+r-1}^{(s)}(u,t)$ are the solutions of the recurrent systems of the partial differential equations

$$\begin{gathered}\frac{\partial H_0^{(s)}}{\partial t}-(\lambda u^2-\beta u)\frac{\partial H^{(s)}}{\partial u}=0,\\ \frac{\partial H_1^{(s)}}{\partial t}-(\lambda u^2-\beta u)\frac{\partial H_1^{(s)}}{\partial u}=\gamma_1^{(s)}(t)\frac{\partial H^{(s)}}{\partial u},\\ \cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial H_{n+r-1}^{(s)}}{\partial t}-(\lambda u^2-\beta u)\frac{\partial H_{n+r-1}^{(s)}}{\partial u}=\gamma_{n+r-1}^{(s)}(t)\frac{\partial H_{n+r-2}^{(s)}}{\partial u},\end{gathered} \tag{46}$$

associated with the boundary conditions $H_0^{(s)}(u,0)=u$, $H_1^{(s)}(u,0) = H_2^{(s)}(u,0) = \ldots = H_{n+r-1}^{(s)}(u,0) = 0$ and corresponding to the particular sequences (44). Using (45), the first term on the right side of formula (43) may be written in the form

$$(47) \qquad \int_0^t \delta(t_0)\left[\frac{\partial G_{n-1,r}(u,t_0)}{\partial u}\right]_{u\sim G_0(u,t-t_0)} dt_0$$

$$= \sum_{s=1}^{\binom{n+r-1}{r}} \int_0^t \delta(t_0)\left[\frac{\partial H_{n+r-1}^{(s)}(u,t_0)}{\partial u}\right]_{u\sim G_0(u,t-t_0)} dt_0 .$$

For every $1 \leqslant s \leqslant \binom{n+r-1}{r}$ the integral

$$(48) \qquad \int_0^t \delta(t_0)\left[\frac{\partial H_{n+r-1}^{(s)}}{\partial u}\right]_{u\sim G_0(u,t-t_0)} dt_0$$

represents the solution of the differential equation

$$(49) \qquad \frac{\partial H_{n+r}^{(s)}}{\partial t} - (\lambda u^2 - \beta u)\frac{\partial H_{n+r}^{(s)}}{\partial u} = \delta(t)\frac{\partial H_{n+r-1}^{(s)}}{\partial u}$$

coupled with the boundary condition $H_{n+r}^{(s)}(u,0)=0$. Thus integral (47) is the sum of the solutions of the $\binom{n+r-1}{r}$ recurrent systems of differential equations consisting of equations (46) and (49). Similarly, it can be shown that also the second term on the right side of formula (43) is the sum of solutions of $\binom{n+r-1}{r-1}$ analogous recurrent systems of partial differential equations. Hence it follows that $G_{n,r}(u,t)$ is the sum of solutions of

$$(50) \qquad \binom{n+r-1}{r} + \binom{n+r-1}{r-1} = \binom{n+r}{r}$$

recurrent systems of partial differential equations. This completes the proof of theorem 2.

From theorem 2 and corollary 1 two corollaries can be easily obtained.

COROLLARY 3. *The partial sums $G_{n,r}(u,t)$ of the expansion of the function $G(u,v,w,t)$ that satisfies equation* (34) *can be presented in the form*

$$(51) \qquad G_{n,r}(u,t)$$

$$= \left(\frac{\lambda}{\beta}\right)^{n+r-1} \frac{\xi_0^{(n,r)}(t) + \left(\frac{\lambda}{\beta}\right)\xi_1^{(n,r)}(t)\,u + \ldots + \left(\frac{\lambda}{\beta}\right)^{n+r+1}\xi_{n+r+1}^{(n,r)}(t)\,u^{n+r+1}}{\left[1-\frac{\lambda}{\beta}(1-\exp\{-\beta t\})\,u\right]^{n+r+1}},$$

where the functions $\xi_0^{(n,r)}(t), \ldots, \xi_{n+r+1}^{(n,r)}(t)$ depend on $\delta(t)$, $\eta(t)$ and β.

Proof. This is a consequence of the fact that $H_{n+r}^{(s)}(u,t)$, where $1 \leqslant s \leqslant \binom{n+r}{r}$, are solutions of systems of partial differential equations analogous to the system consisting of (46) and (49) corresponding to different sequences $\gamma_1^{(s)}(t), \ldots, \gamma_{n+r}^{(s)}(t)$ and can be presented either in the form

$$(52) \quad H_{n+r}^{(s)}(u, t_{n+r}) = \int_0^{t_{n+r}} \gamma_{n+r}^{(s)}(t_{n+r-1}) \int_0^{t_{n+r-1}} \gamma_{n+r-1}^{(s)}(t_{n+r-2}) \ldots \ldots \int_0^{t_1} \gamma_1^{(s)}(t_0) W_{n+r}(u, t_{n+r}, \ldots, t_0)\, dt_0 \ldots dt_{n+r-1},$$

where $W_{n+r}(u, t_{n+r}, t_{n+r-2}, \ldots, t_0)$ is a sum of expressions of the form

$$(53) \quad A_{n+r}^{(s)} \left[\frac{G_0(u, t_{n+r})}{G_0(u, t_{n+r}-t_0)}\right]^{\varepsilon_0} \left[\frac{G_0(u, t_{n+r})}{G_0(u, t_{n+r}-t_1)}\right]^{\varepsilon_1} \ldots \left[\frac{G_0(u, t_{n+r})}{G_0(u, t_{n+r}-t_{n+r-1})}\right]^{\varepsilon_{n+r-1}}$$

where the ε_i equal $0, 1$ or 2, while

$$\sum_{i=0}^{n+r-1} \varepsilon_i = n+r+1$$

and $A_{n+r}^{(s)}$ does not depend on u, or in the form

$$(54) \quad H_{n+r}^{(s)}(u, t) = \left(\frac{\lambda}{\beta}\right)^{n+r-1} \frac{\xi_0^{(n,r,s)}(t) + \left(\frac{\lambda}{\beta}\right) \xi_1^{(n,r,s)}(t)\, u + \ldots + \left(\frac{\lambda}{\beta}\right)^{n+r+1} \xi_{n+r+1}^{(n,r,s)}(t)\, u^{n+r+1}}{\left[1 - \frac{\lambda}{\beta}(1 - \exp\{-\beta t\})\, u\right]^{n+r+1}},$$

where the functions $\xi_0^{(n,r,s)}(t), \ldots, \xi_{n+r+1}^{(n,r,s)}(t)$ depend on $\delta(t)$, $\eta(t)$ and β. The proof of (52) and (54) proceeds in an analogous manner as the one of formulae (22) and (7). Hence (51) follows, because the sum of expressions (54) has the form of (51).

Using (51) and (29) we obtain for $p_{m,n,r,}(t)$ the formula

$$(55) \quad p_{m,n,r}(t) = \left(\frac{\lambda}{\beta}\right)^{m+n+r-1} (1 - e^{-\beta t})^m \sum_{i=0}^{j} \binom{m+n+r+i}{n+r} \frac{\xi_i^{(n,r)}(t)}{(1 - e^{-\beta t})^i},$$

where $j = \min(m, n+r+1)$. Note that in this case the $n+r+2$ functions $\xi_0^{(n,r)}(t), \ldots, \xi_{n+r+1}^{(n,r)}(t)$ appear in the formula for $p_{m,n,r}(t)$. Also, if $p_{0,n,r}(t), p_{1,n,r}(t), \ldots, p_{n+r+1,n,r}(t)$ are known, the functions $\xi_0^{(n,r)}(t), \ldots, \xi_{n+r+1}^{(n,r)}(t)$ can be evaluated from (55).

COROLLARY 4. *If $\delta(t)$ and $\eta(t)$ are constant, then the partial sums $G_{n,r}(u, t)$ of the expansion of the function $G(u, v, w, t)$ that satisfies equation* (34) *can be expressed by the partial sums $G_{n+r}(u, t)$,*

$$G_{n,r}(u, t) = \binom{n+r}{r} \delta^n \eta^r G_{n+r}(u, t), \tag{56}$$

of the expansion of the function $G(u, v, t)$ that satisfies equation (1) *with* $\gamma(t) = 1$.

Proof. In view of the assumption of corollary 4 formula (52) reduces to

$$H_{n+r}^{(s)}(u, t_{n+r}) = \delta^n \eta^r \int_0^{t_{n+r}} \int_0^{t_{n+r-1}} \ldots \int_0^{t_0} W_{n+r}(u, t_{n+r}, \ldots, t_0) dt_0 \ldots dt_{n+r-1}. \tag{57}$$

Finally, using (16) we get

$$H_{n+r}^{(s)}(u, t_{n+r}) = \delta^n \eta^r G_{n+r}(u, t_{n+r})$$

and formula (56) follows from theorem 2.

From (56) interesting formulae

$$G_{n,r}(u, t) = \binom{n+r}{r}\left(\frac{\delta}{\beta}\right)^n \left(\frac{\eta}{\beta}\right)^r \left(\frac{\lambda}{\beta}\right)^{n+r-1} \times \tag{58}$$

$$\times \frac{\varphi_0^{(n+r)}(\beta t) + \left(\frac{\lambda}{\beta}\right) \varphi_1^{(n+r)}(\beta t) u + \ldots + \left(\frac{\lambda}{\beta}\right)^{n+r+1} \varphi_{n+r+1}^{(n+r)}(\beta t) u^{n+r+1}}{\left[1 - \frac{\lambda}{\beta}(1 - \exp\{-\beta t\}) u\right]^{n+r+1}}$$

and

$$p_{m,n,r}(t) \tag{59}$$

$$= \binom{n+r}{r}\left(\frac{\delta}{\beta}\right)^n \left(\frac{\eta}{\beta}\right)^r \left(\frac{\lambda}{\beta}\right)^{m+n+r-1} (1 - e^{-\beta t})^m \sum_{i=1}^{j} \binom{m+n+r-i}{n+r} \frac{\varphi_i^{(n+r)}(\beta t)}{(1 - e^{-\beta t})^i}$$

can be obtained, where $j = \min(m, n+r+1)$.

It is of interest that both coefficients (33) and (59) can be expressed by the same functions $\varphi_0^{(k)}(t), \ldots, \varphi_{k+1}^{(k)}(t)$, where $k = 0, 1, 2, \ldots$

3. Applications. The main result of the previous section is formula (31) for the coefficients of the expansion of the function satisfying equation (1). If differential equations (1) and (34) represent equations for the generating functions, then (31) and (55) provide the formulae for the corresponding probabilities describing the stochastic models. Three examples where these formulae are applicable are given in this section. The first two deal with the birth and death processes with constant rates.

The last application is to the joint distribution arising in the stochastic model of carcinogenesis of Neyman [3].

Example 1. If $v = 1$, $\gamma(t) = \mu$ and $\beta = \lambda + \mu$ are substituded in differential equation (1), then

$$G(u, t) = \mu B(t) + [1 - \lambda B(t)][1 - \mu B(t)] \sum_{m=1}^{\infty} [\lambda B(t)]^{m-1} u^m, \tag{60}$$

where

$$B(t) = \frac{1 - \exp\{t(\lambda - \mu)\}}{\mu - \lambda \exp\{t(\lambda - \mu)\}}. \tag{61}$$

This is the generating function of a pure birth and death process with constant rates λ and μ. If, in addition to the number of live cells, say $x(t)$, the number of cells, say $y(t)$, dying before time t, is taken into account, the differential equation for the generating function $G(u, v, t)$ is

$$\frac{\partial G}{\partial t} - [\lambda u^2 - (\lambda + \mu) u + \mu v] \frac{\partial G}{\partial u} = 0, \tag{62}$$

and let $G(u, v, 0) = u$.

Equation (62) is obtained in the usual way by letting

$$p_{m,n}(t) = P\{x(t) = m, y(t) = n\}. \tag{63}$$

Now it is a simple matter to show that

$$p'_{m,n}(t) = \lambda(m-1) p_{m-1,n}(t) + \mu(m+1) p_{m+1,n-1}(t) - (\lambda + \mu) m p_{m,n}(t). \tag{64}$$

Hence equation (62) for the generating function

$$G(u, v, t) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} p_{m,n}(t) u^m v^n \tag{65}$$

follows easily.

Under the assumption that λ and μ are constant the solution

$$G(u, v, t) = \frac{u_1(u - u_2) - u_2(u - u_1) \exp\{-\lambda(u_2 - u_1)t\}}{u - u_2 - (u - u_1) \exp\{-\lambda(u_2 - u_1)t\}} \tag{66}$$

of (62) is easily obtained by the well-known methods, where

$$\begin{aligned} u_1 &= \frac{\lambda + \mu + [(\lambda - \mu)^2 + 4\lambda\mu(1 - v)]^{1/2}}{2\lambda}, \\ u_2 &= \frac{\lambda + \mu - [(\lambda - \mu)^2 + 4\lambda\mu(1 - v)]^{1/2}}{2\lambda}. \end{aligned} \tag{67}$$

However, the solution is complicated, so that it is tedious to find the coefficients of the expansion of the function $G(u, v, t)$. Hence formula (33) is proposed to evaluate the probabilities $p_{m,n}(t)$.

Example 2. A birth and death process with mutation will be considered next. Let $x(t)$ and $y(t)$ be defined as previously, and let $z(t)$ be the number of mutations in the time interval from 0 to t. If

$$p_{m,n,r}(t) = P\{x(t) = m, y(t) = n, z(t) = r\}, \tag{68}$$

then

$$\begin{aligned} p'_{m,n,r}(t) = {} & \lambda(m-1)p_{m-1,n,r}(t) + \mu(m+1)p_{m+1,n-1,r}(t) + \\ & + \nu(m+1)p_{m+1,n,r-1}(t) - (\lambda+\mu+\nu)mp_{m,n,r}(t), \end{aligned} \tag{69}$$

and

$$\frac{\partial G}{\partial t} - [\lambda u^2 - (\lambda+\mu+\nu)u + \mu v + \nu w]\frac{\partial G}{\partial u} = 0. \tag{70}$$

Suppose that $G(u, v, w, 0) = u$. Here λ, μ, ν are the rates of birth, death and mutation, respectively, and $G(u, v, w, t)$ is the corresponding generating function.

The formulae for $p_{m,n,r}(t)$ can be obtained from (59).

Example 3. The stochastic model of carcinogenesis of Neyman is a two-stage birth and death process. It is convenient to call the two different kinds of cells considered in the model grey cells and black cells. The grey cells multiply and die and, moreover, they are exposed to a risk of mutation. The result of the mutation is a black cell. Let λ be the rate of birth, μ of death and ν of mutation. The black cells can only multiply and die. For the black cells let L be the rate of birth and M of death. Every grey cell and every black cell that arises as the result of a mutation of a grey cell develops into a clone according to a birth and death process. For simplicity it will be assumed that initially there is only one grey cell, the ancestor of all future grey and black cells. This assumption is not made in the model of Neyman, but it puts no essential restrictions on the proposed method to evalute the final distribution. Assume that the probability that a black clone of m cells is visible is Φ_m, with $0 \leqslant \Phi_m \leqslant 1$ and $\Phi_0 = 0$. The subject of the study is the joint distribution of the number of grey cells and the number of visible black clones at time T.

The basic differential equation for the generating function has been deduced by Neyman [3] as follows. Let $X(t)$ be the number of grey live cells at time t and $Y(t)$ the number of those black clones that have been generated by mutation before time t and will be visible at time $T > t$.

Let

$$q_{m,n}(t) = \{X(t) = m, Y(t) = n\}. \tag{71}$$

Using these assumptions and formulae (60) and (61) the probability that a clone of black cells generated by a mutation at time t will be visible at time T is equal to

$$(72)\quad \Pi(T-t) = [1-LB(T-t)][1-MB(T-t)]\sum_{m=1}^{\infty}[LB(T-t)]^{m-1}\Phi_m.$$

Now, by the usual argument, the system of differential equations corresponding to the joint distribution of the number of grey cells and the number of visible black clones at time T will be found to be

$$(73)\quad q'_{m,n}(t) = \lambda(m-1)q_{m-1,n}(t)+\Pi(T-t)\nu(m+1)q_{m+1,n-1}(t)+ \\ +\big[\mu+\big(1-\Pi(T-t)\big)\nu\big](m+1)q_{m+1,n}(t)-(\lambda+\mu+\nu)mq_{m,n}(t).$$

Hence

$$(74)\quad \frac{\partial G}{\partial t}-\big\{\lambda u^2-(\lambda+\mu+\nu)u+\Pi(T-t)\nu v+\big[\mu+\big(1-\Pi(T-t)\big)\nu\big]\big\}\frac{\partial G}{\partial u} = 0,$$

where $G(u, v, w, t)$ is the generating function associated with the considered joint distribution. The boundary condition is $G(u, v, w, 0) = u$. This is the basic equation obtained by Neyman.

We now proceed to determine the probabilities $q_{m,n}(t)$. If the notation

$$(75)\quad \delta(t) = \Pi(T-t)\nu,$$

$$(76)\quad \eta(t) = \mu+[1-\Pi(T-t)]\nu,$$

$$(77)\quad \beta = \lambda+\mu+\nu$$

is introduced, the differential equation (74) becomes

$$(78)\quad \frac{\partial G}{\partial t}-\big(\lambda u^2-\beta u+\delta(t)v+\eta(t)\big)\frac{\partial G}{\partial u} = 0.$$

However, relation (78) is the differential equation (34) with $w = 1$ considered in section 2. It is now clear that

$$(79)\quad q_{m,n}(t) = \sum_{r=0}^{\infty}p_{m,n,r}(t),$$

where $p_{m,n,r}(t)$ are the coefficients of the expansion of the function $G(u, v, w, t)$ satisfying equation (34), while $\delta(t)$, $\eta(t)$ and β are given by (75), (76) and (77).

The probabilities $p_{m,n,r}(t)$ can be computed, as was shown in section 2. Then approximate values of the $q_{m,n}(t)$ can be evaluated from the sum (79).

REFERENCES

[1] M. S. Bartlett, *Stochastic processes* (notes of a course given at the University of North Carolina), 1946.

[2] D. G. Kendall, *Birth- and death-processes, and the theory of carcinogenesis*, Biometrika 47(1960), p. 13-22.

[3] J. Neyman, *A two-step mutation theory of carcinogenesis*, Bulletin of the International Statistical Institute 38 (1961), p. 123-135.

[4] M. J. Polissar and M. B. Shimkin, *A quantitative interpretation of the distribution of induced pulmonary tumors in mice*, Journal of the National Cancer Institute 15 (1954), p. 377-403.

Reçu par la Rédaction le 24. 9. 1964