AGE DISTRIBUTIONS IN DIVIDING POPULATIONS

GROVE C. NOONEY

From the Lawrence Radiation Laboratory, University of California, Berkeley

ABSTRACT We show that the age distribution tends to a limit for each population of cells that die or divide according to continuous age-dependent schedules. This limiting distribution is independent of the initial age distribution. Explicit formulas are given for the limiting age distributions and for all stationary age distributions. Nonstationary behavior periodic in time is impossible.

INTRODUCTION

We consider from a deterministic viewpoint the growth of populations of cells that die or divide according to continuous age-dependent schedules. We show that the age distribution of any such population tends to a limit that is independent of the initial age distribution. Further, we exhibit explicit formulas for the limiting distributions and, consequently, for the stationary age distribution of each population. Convergence to a limiting age distribution differs from the periodic behavior derived by other authors (Trucco, 1965 *b*, who also gives additional references). They described periodic behavior only under the unnatural hypothesis that all cells considered divide at precisely the same age. As we show here, even slight deviations from such a rule, as certainly occur in natural populations, will destroy periodicity.

Our conclusions make more definite the usually vague invocations of "steady-state kinetics" in experimental studies of population growth or metabolism, for if age governs growth or metabolism, then the attainment of a stationary age distribution permits the application of steady-state methods. Our conclusions may even allow the use of naturally growing cell cultures in experiments formerly thought to require synchronization, for a close approximation to a known stationary age distribution may serve instead of a close approximation to a known single age.

More precisely, we consider a continuous finite population of cells, measuring the age of each cell from its time of birth. Let $n(a, t)$ be the number of cells in the population that are of age a at time t. By continuous population, we mean that $n(a, t)$ is a continuous function of a for each $t \geqslant 0$. This is a departure from actual populations of cells, where for each t the $n(a, t)$ can be nonzero only at a finite number of distinct ages a. For large populations, the smoothing introduced by continuity is thought to be negligible, particularly since the determination of age by experimental means is necessarily approximate. Let $N(a, t) = \int_0^a n(s, t)\, ds$. Then

$N(a, t)$ represents the number of cells of age not exceeding a at time t. By finite population, we mean that $N(\infty, t) = \lim_{a\to\infty} N(a, t)$ exists. We assume that the partial derivatives N_t and N_{at} exist and are continuous. For each population, we define the age distribution D by $D(a, t) = N(a, t)/N(\infty, t)$, with the convention that $D(a, t) = 1$ when $N(\infty, t) = 0$.

Now suppose that $P(a)$ is the proportion of cells of age zero that would divide before age a if no deaths were to occur, and suppose that $Q(a)$ is the proportion of cells of age zero that would die before age a if no divisions were to occur. Division means replacement by two replicas of age zero; death means removal from the population. We assume that $Q(0) = P(0) = 0$, that $\lim_{a\to\infty} P(a) = \lim_{a\to\infty} Q(a) = 1$, and that P and Q are continuously differentiable. Note that discontinuous P or Q are excluded. We set $I = 2\int_0^\infty [1 - Q(s)]P'(s)\,ds$, the average number of first generation descendants per cell.

If $t = 0$ for some arbitrary initial time, it is clear that $N(a, t)$ depends on $N(a, 0)$ for all $a > 0$. As we shall show, however, the influence of $N(a, 0)$ is negligible for large t. The main result to be demonstrated is that $\lim_{t\to\infty} D(a, t)$ exists and is independent of $N(a, 0)$. Indeed, we shall find explicitly the asymptotic behavior of $N(a, t)$ for large t: when $I < 1$, $N(a, t)$ tends to zero for each a; when $I > 1$, $N(a, t)$ increases exponentially in t; when $I = 1$, $N(a, t)$ tends to a function of a alone. The condition $I = 1$ deserves particular and further attention for moderate values of t, but we postpone its detailed consideration.

The demonstration proceeds through three parts: the establishment of a basic equation for $N(a, t)$; the derivation of a renewal equation, the solution of which governs $N(a, t)$; and the analysis of that renewal equation (Asymptotic Behavior). We have relegated details of the asymptotic analysis to the Appendix. A few remarks conclude the paper. Consideration of renewal equations and their asymptotic behavior is not new in the context of populations of cells (Bellman and Cooke, 1963; Harris, 1963; Trucco, 1965 *b*), but the main results of our consideration is new.

A BASIC EQUATION

Let $p(s) = P'(s)[1 - P(s)]^{-1}$ and $q(s) = Q'(s)[1 - Q(s)]^{-1}$, and let o generically denote a function such that $\lim_{x\to 0} o(x)/x = 0$. Then the proportions of cells of age s which divide or die in the age interval $(s, s + \delta t)$ are given respectively by $p(s)\delta t + o(\delta t)$ and $q(s)\delta t + o(\delta t)$. The proportion of cells of age s which neither die nor divide in the age interval $(s, s + \delta t)$ is given by $[1 - p(s)\delta t][1 - q(s)\delta t] + o(\delta t)$. Our stated assumptions lead to the equation

$$N(a, t + \delta t) = \int_{s=0}^{a-\delta t} N_a(s, t)[1 - p(s)\delta t][1 - q(s)\delta t]\,ds + 2\int_{s=0}^{\infty} N_a(s, t)p(s)\delta t\,ds + o(\delta t). \tag{1}$$

The integrals on the right side of equation (1) represent the respective contributions to $N(a, t + \delta t)$ of cells that neither die nor divide in the time interval $(t, t + \delta t)$ and of cells that do divide in that interval. Rearranging equation (1) and carrying out the integration involving only N_a, we find

$$[N(a, t + \delta t) - N(a - \delta t, t)](\delta t)^{-1} = -\int_{s=0}^{a-\delta t} N_a(s, t)[p(s) + q(s)]\, ds + 2\int_{s=0}^{\infty} N_a(s, t)\ p(s)\ ds + \frac{o(\delta t)}{\delta t}.$$

Letting δt tend to zero, we obtain the sought equation

$$\frac{dN(a, t)}{dt} = 2\int_{s=0}^{\infty} N_a(s, t)p(s)\ ds - \int_{s=0}^{a} N_a(s, t)[p(s) + q(s)]\, ds, \qquad (2)$$

where d/dt stands for $\partial/\partial a + \partial/\partial t$. As we shall show, the function $N(a, t)$ is uniquely determined by equation (2) and the function $N(a, 0)$, with $a \geqslant 0$.

A RENEWAL EQUATION

By differentiation of equation (2) with respect to a, we obtain the partial differential equation mentioned by several authors:

$$\frac{dn(a, t)}{dt} = -[p(a) + q(a)]n(a, t). \qquad (3)$$

[Trucco, 1965 *a* and 1965 *b* (who gives additional references)]. As indicated elsewhere (Trucco, 1965 *a*), this equation together with $N(a, 0)$ or $n(a, 0)$ for $a \geqslant 0$ is not sufficient to determine uniquely $n(a, t)$ or $N(a, t)$. This is to be expected since equation (3) reflects the process of attrition but not the process of replication.

By virtue of our knowledge of the general solution of equation (3) (Kamke, 1944, p. 141), satisfied by $n(a, t) = N_a(a, t)$, we know that $N(a, t)$ must have the form

$$N(a, t) = \int_{s=0}^{a} N_a(s - t, 0) \exp[R(s - t) - R(s)]\, ds, \qquad (4)$$

where

$$R(a) = \int_{s=0}^{a} [p(s) + q(s)]\, ds \qquad (5)$$

and where $N_a(s - t, 0)$ is given only for $s \geqslant t$ by $N(a, 0)$ with $a \geqslant 0$. For $s < t$, the unknown function $N_a(s - t, 0) \exp R(s - t)$ reflects the replicative process and represents entry of cells of age zero into the population at time $t - s$ (Trucco, 1965 *b*). The relation (5) may be written more explicitly:

$$R(a) = -\log \{[1 - P(a)][1 - Q(a)]\}. \tag{6}$$

We regard $N_a(s, 0) \exp R(s)$ for $s < 0$ simply as an auxiliary function, and to determine it we set $y(-s)$ equal to $N_a(s, 0) \exp R(s)$ and write equation (4) as

$$N(a, t) = \int_{s=0}^{a} y(t - s)[1 - P(s)][1 - Q(s)]\, ds. \tag{7}$$

The replacement of $N(a, t)$ in equation (2) by this expression and an integration by parts lead to the integral equation for y,

$$y(t) = 2 \int_{s=0}^{\infty} y(t - s)[1 - Q(s)]P'(s)\, ds. \tag{8}$$

Specification of $N(a, 0)$, with $a \geqslant 0$, determines $y(t - s)$ for $t \leqslant s$, and we may therefore write equation (8) in the form of a classical renewal equation,

$$y(t) = 2 \int_{s=0}^{t} y(t - s)[1 - Q(s)]P'(s)\, ds + f(t), \tag{9}$$

where

$$f(t) = 2 \int_{s=t}^{\infty} N_a(s - t, 0) \exp [R(s - t) - R(s)]p(s)\, ds \tag{10}$$

is a known function, given $N_a(a, 0)$, $P(a)$, and $Q(a)$ for $a \geqslant 0$.

So far, we have proceeded without regard for differentiability and integrability requirements. If $y(t)$ is to be continuous at $t = 0$, then the definition, $y(-s) = N_a(s, 0) \exp R(s)$, together with equations (9) and (10), yields the compatibility condition $N_a(0, 0) = f(0)$ or

$$N_a(0, 0) = 2 \int_{0}^{\infty} N_a(s, 0)p(s)\, ds. \tag{11}$$

Let us assume for the moment that this compatibility condition holds. If equation (2) is to have meaning for $t = 0$, we must assume the existence of

$$\int_{s=0}^{\infty} N_a(s, 0)p(s)\, ds \quad \text{and} \quad \int_{s=0}^{\infty} N_a(s, 0)q(s)\, ds.$$

It is sufficient to require further the existence and continuity of $N_{aa}(s, 0)$ for $s \geqslant 0$ and the existence of $f(0)$ and $f'(0)$, in addition to the restrictions already placed on P and Q. For then the function f is continuous and bounded, entailing the existence of a unique, continuous solution y of the renewal equation (9) (Bellman and Cooke, 1963). Since, further, f is continuously differentiable and all other functions appearing in equation (9) are continuous, it follows that y is continuously differentiable. Therefore, the function N defined by equation (7) is differentiable in each of its variables, and the operations required for the formation of equation (2), as well as those performed in the derivation of equation (9), are all permitted.

Equations (7) and (9) allow the calculation of $N(a, t)$ for any desired (a, t), but let us turn to the behavior of $N(a, t)$ for large t.

ASYMPTOTIC BEHAVIOR

The conduct of $N(a, t)$ for large t is determined through equation (7) by the asymptotic behavior of y. The latter is greatly influenced by the kernel of the renewal equation (9). We may distinguish three cases, according as

$$I = 2 \int_0^\infty [1 - Q(s)]P'(s)\, ds$$

is (i) less than, (ii) greater than, or (iii) equal to unity. In case (ii), there exists a unique constant $c > 0$ such that

$$2 \int_0^\infty e^{-cs}[1 - Q(s)]P'(s)\, ds = 1, \tag{12}$$

for considered as a function of c, the left side of this equation is continuous and monotonically decreasing, exceeds unity for $c = 0$, and is arbitrarily small for c sufficiently large. Under certain conditions, the following behavior of y is established for the three cases: (i) $\lim_{t\to\infty} y(t) = 0$, (ii) $\lim_{t\to\infty} y(t)e^{-ct} = K_1$, and (iii) $\lim_{t\to\infty} y(t) = K_2$ for suitable positive constants, K_1 and K_2. Proof of this behavior detailed conditions, and explicit K_1 and K_2 are given in the Appendix.

The consequences for $N(a, t)$ of the various asymptotic behaviors of y are easily assessed through equation (7). Indeed, we find

$$\lim_{t\to\infty} N(a, t) = 0 \qquad \text{for } I < 1,$$

$$\lim_{t\to\infty} N(a, t)e^{-ct} = K_1 \int_{s=0}^{a} e^{-cs}[1 - P(s)][1 - Q(s)]\, ds \qquad \text{for } I > 1,$$

and

$$\lim_{t\to\infty} N(a, t) = K_2 \int_{s=0}^{a} [1 - P(s)][1 - Q(s)]\, ds \qquad \text{for } I = 1,$$

where the constant $c > 0$ is given implicitly by equation (12). It follows that $D(a, t) = N(a, t)/N(\infty, t)$ tends to a limiting distribution in each case and that the limiting distribution is independent of the initial distribution.

If $N(a, 0)$ is taken as proportional to the appropriate limiting distribution, then $D(a, t)$ will be stationary or independent of t. That these are the only stationary $D(a, t)$ follows from the necessary convergence to a limiting distribution.

REMARKS

The form of stationary $D(a, t)$ could have been obtained in other ways. For instance, if $D(a, t) = N(a, t)/N(\infty, t)$ is stationary, then $N(a, t)$ is the product of two functions of a single variable each. We may then treat equation (2) by the method of separation of variables and obtain $N(a, t)$ by solving two ordinary differential equations. The point of the present work is, however, that all age distributions converge to certain limiting distributions. That these limiting distributions are the stationary distributions is incidental to our work but provides considerable insight into the nature of the stationaries.

Regarding P and Q as being fixed functions, we impose several conditions on $N(a, 0)$ and its derivatives, usually indirectly by way of restrictions on the function f defined by equation (10). Apart from the smoothness required by our formal manipulations, these restrictions ensure only that $n(a, 0) = N_a(a, 0)$ tends rapidly enough to zero as $P(a)$ or $Q(a)$ tends to unity. Thus, we may say that we require a sufficiently youthful population given by $N(a, 0)$. It is easily seen that naturally growing populations have precisely this quality: few cells survive past an age a at which $P(a)$ or $Q(a)$ is almost unity.

The compatibility condition (11) is also a restriction to a natural initial population, ensuring that the initial population arose from the process defined by P and Q. It is not an essential condition, however. If the compatibility condition does not hold, we determine y from the renewal equation (9) and then redefine $y(0) = N_a(0, 0)$. The resulting $N(a, t)$, given by equation (7), then has a derivative saltus at $a = t$. Oevertheless, our conclusions for $N(a, t)$ remain valid for $a < t$. Under the condition of a finite life span, the limiting age distributions shown are valid also.

Our results can be extended slightly to include the case in which each dividing cell gives rise to k daughters. The conclusions remain valid if we replace I by $kI/2$ and correspondingly modify the definitions of c, m, and f.

The stochastic treatment of questions of age distributions has received attention in the case of age-dependent birth and death probabilities (Kendall, 1949), but most questions remain open. Our work may be interpreted as a contribution to a stochastic theory, for if P and Q are regarded as distribution functions of random variables, then our results pertain to the mean value of the random variable $N(a, t)$. Our assertions about convergence to limiting distributions are less meaningful without the evaluation of other statistics such as the variance of $N(a, t)$. Since actual populations of cells exhibit stochastic behavior, practical applications make this difficult evaluation of immediate import.

APPENDIX

Here we determine the asymptotic behavior of the solution y of the renewal equation

$$y(t) = 2\int_{s=0}^{t} y(t - s)[1 - Q(s)]P'(s)\,ds + f(t).$$

We set $I = 2\int_0^\infty [1 - Q(s)]P'(s)\,ds$ and define

$$F(x) = \int_0^\infty e^{-xs}f(s)\,ds, \qquad D(x) = \int_0^\infty \left|\frac{d}{ds}e^{xs}f(s)\right| ds$$

and

$$m(x) = 2\int_0^\infty se^{-xs}[1 - Q(s)]P'(s)\,ds.$$

Distinguishing three cases: (i) $I < 1$, (ii) $I > 1$, and (iii) $I = 1$, we show the corresponding asymptotic behavior of y: (i) $\lim_{t\to\infty} y(t) = 0$, (ii) $\lim_{t\to\infty} y(t)e^{-ct} = F(c)/m(c)$, where c is given implicitly by equation (12), and (iii) $\lim_{t\to\infty} y(t) = F(0)/m(0)$.

Proof of this behavior is in several theorems which we cite by number from Bellman and Cooke (1963). Convenient additional assumptions for this proof for the respective cases are the existence of: (i) $D(0)$; (ii) $D(0)$, $F(c)$, and $m(c)$, where c is given by equation (12); and (iii) $F(0)$, $m(0)$, and a $b > 0$ for which $\lim_{t\to\infty} e^{bt}f(t) = 0$ and for which $D(b)$ exists. First, we may represent y in the form (theorem 7.6)

$$y(t) = f(0)u(t) + \int_0^t u(t - s)f'(s)\,ds, \tag{13}$$

where u is the continuous solution of the auxiliary renewal equation

$$u(t) = 2\int_0^t u(t - s)[1 - Q(s)]P'(s)\,ds + 1.$$

Applied to this auxiliary equation, theorems 7.14 and 7.11 yield the results: $\lim_{t\to\infty} u(t) = 1/(1 - I)$ in case (i), and $\lim_{t\to\infty} u(t)e^{-ct} = 1/cm(c)$ in case (ii). Equation (13) then permits the stated conclusions for y in cases (i) and (ii).

Case (iii) can be reduced to case (ii) by considering $z(t) = e^{bt}y(t)$, where b is any number satisfying our assumptions for case (iii). We easily obtain the renewal equation

$$z(t) = \int_0^t z(t - s)e^{bs}[1 - Q(s)]P'(s)\,ds + e^{bt}f(t),$$

which possesses a case (ii) kernel. Since $D(b)$ exists, $z(t)$ has a representation of the form of equation (13), and we conclude from the behavior of the solution of its auxiliary equation that $\lim_{t\to\infty} y(t) = \lim_{t\to\infty} e^{-bt}z(t) = F(0)/m(0)$.

Relaxation of our assumptions is possible under the condition of a finite life span for the cells considered. Namely, if there is a number A such that $P(a) = Q(a) = 1$ when $a > A$, then we need not assume the existence of $F(0)$, $F(c)$, $m(0)$, $m(c)$, or $D(0)$.

This work was supported by the U. S. Atomic Energy Commission.

Received for publication 7 October 1966.

BIBLIOGRAPHY

Bellman, Richard, and Kenneth L. Cooke. 1963. Differential-Difference Equations. Academic Press, Inc., New York. Chapter 7.

HARRIS, THEODORE E. 1963. The Theory of Branching Processes. Springer Verlag, Berlin-Göttingen-Heidelberg.

KAMKE, E. 1944. Differentialgleichungen, Lösungsmethoden und Lösungen. Becker & Erler Kom.-Ges., 2. Leipzig. Akademische Verlagsgesellschaft.

KENDALL, DAVID G. 1949. *J. Roy. Statist. Soc. London, Series B.* **11:** 230.

TRUCCO, E. 1965*a*. *Bull. Math. Biophys.* **27:** 285.

TRUCCO, E. 1965 *b*. *Bull. Math. Biophys.* **27:** 449.