On Some Modes of Population Growth Leading to R. A. Fisher's Logarithmic Series
Distribution

Author(s): David G. Kendall

# ON SOME MODES OF POPULATION GROWTH LEADING TO R. A. FISHER'S LOGARITHMIC SERIES DISTRIBUTION

### By DAVID G. KENDALL, M.A.
*Magdalen College, Oxford*

1. R. A. Fisher (1943) in a co-operative study written with A. S. Corbet and C. B. Williams has developed a mathematical theory which describes with some success the relative numbers of animals of different species obtained when sampling at random from a heterogeneous population. This problem was first considered in relation to (i) Corbet's work on the distribution of butterflies in the Malay Peninsula, and (ii) the numbers of moths of different species caught in a light-trap over a given period of time (Williams's data). Fisher began by assuming that for a particular species the number of individuals caught in time $t$ would be distributed as a Poisson variable of expectation $\omega t$, where $\omega$ may be called the *intrinsic abundance* of the species. He suggested that $\omega$ might be distributed in the Eulerian (or $\chi^2$) form

$$\frac{1}{\Gamma(k)}\left(\frac{k}{\Omega}\right)^k e^{-k\omega/\Omega}\omega^{k-1}\,d\omega \quad (0<\omega<\infty), \tag{1}$$

where $\Omega$ is the mean value of $\omega$ and $k$ is a constant parameter, and showed that the actual number caught would then follow a negative binomial distribution with index $k$[*]. In fitting such a distribution to Corbet's data he obtained very small values of $k$, and this suggested that it might be worth while examining what would happen if $\Omega$ and $k$ were allowed to tend to zero in a constant ratio. In this way Fisher found that if a species were known to have been caught, it would be represented in the catch by exactly $n$ individuals with a probability

$$\frac{x^n}{ny} \quad (n = 1, 2, 3, \ldots), \tag{2}$$

where[†]
$$y = \ln\frac{1}{1-x}, \quad x = \frac{at}{1+at} \quad \text{and} \quad a = \lim \Omega/k.$$

The success of this 'logarithmic series distribution' in graduating the entomological data of Corbet, Williams and others implies that in the populations concerned the distribution of intrinsic abundance must be (for $\omega$ not too small) effectively of the form

$$Ae^{-\omega/a}\,d\omega/\omega \quad (A = \text{constant}). \tag{3}$$

(This cannot of course be true for *all* $\omega$, for then the integral of total probability would not converge.) It will be noticed that the distribution (3) of intrinsic abundance is itself the continuous analogue of the logarithmic series. The success of (2) in describing the relative numbers of individuals caught is thus a challenge to biologists to provide a theoretical interpretation for (3).

In this connexion it is worth noting that if one is concerned with a population containing only a finite number ($Z$, say) of species, then the continuous distribution (3) can be replaced by a logarithmic series, and results similar to those of Fisher follow as before. Thus, suppose that the actual number $\nu$ of individuals by which a particular species is represented *in the whole population* is distributed in the discrete form

$$\frac{X^\nu}{\nu Y} \quad (\nu = 1, 2, 3, \ldots), \tag{4}$$

---

[*] This step in the argument is of course equivalent to that taken by M. Greenwood & G. U. Yule (1920) in another context.    [†] I write $\ln z$ for the natural logarithm of $z$.

where $Y = -\ln(1-X)$, and let $p = 1 - e^{-\gamma t}$ be the chance that an individual will be caught in an exposure of duration $t$. Then the chance that the species will have $n = 0, 1, 2, 3, \ldots$ representatives *in the sample* is given by*

$$P_0 = 1 - \frac{y}{Y} \quad \text{and} \quad P_n = (1-P_0)\frac{x^n}{ny} \quad (n>0), \tag{5}$$

i.e. a logarithmic distribution with a zero term added. Here

$$x = 1 - e^{-y} \quad \text{and} \quad e^y - 1 = (e^Y - 1)(1 - e^{-\gamma t}), \tag{6}$$

and so $y \leqslant Y$ for all $t$; for very small exposures $p$ will be small and then

$$y \simeq \gamma t(e^Y - 1),$$

while for very long exposures $p$ will be nearly equal to unity and then

$$y \simeq Y.$$

The expected values of $S$ and $N$ (the number of species and the number of individuals in the catch) are

$$\bar{S} = Z(1-P_0) = yZ/Y \quad \text{and} \quad \bar{N} = (e^y - 1)Z/Y.$$

Thus as $t \to \infty$, $\bar{S} \to Z$ and $\bar{N} \to (e^Y - 1)Z/Y$, while for *all* values of $t$

$$\bar{S} = \alpha \ln(1 + \bar{N}/\alpha), \tag{7}$$

where $\alpha = Z/Y$ is a constant independent of the time of exposure and corresponding to Fisher's 'index of diversity'. Formula (7) is, in fact, identical with the well-known result due to Fisher (1943), although the derivation given here proceeds from somewhat different assumptions.

Williams (1944 $a, b$) has shown that the logarithmic series (2) can also be applied to a great variety of other biological problems, in which the integer $n$ is variously the number of species per genus, the number of genera per subfamily, the number of parasites per host, and even the number of research papers per biologist (published in a particular year). It is hard to believe that a single mechanism will be found to explain the relevance of the logarithmic series to *all* these problems, and it seems therefore well worth while to record any theoretical models which may be found to lead mathematically to this distribution. In the remainder of this paper I shall describe a number of discontinuous Markoff processes which lead to distributions of negative binomial and logarithmic series form, in the hope that some of these may be found to be of biological significance.

2. The stochastic processes to be considered here will for convenience be described in relation to the growth of a hypothetical population of organisms, whose numbers fluctuate with the incidence of mortality, reproduction (by binary fission) and immigration from the outside world. Let $n$ be the size of the population at time $t$; then $n(t)$ is a random function which develops in the following manner:

(i) If $n > 0$, the only possible transitions in an element of time $dt$ are from $n$ to $n-1$, $n$ or $n+1$, and the transition probabilities are

$$n \to \begin{cases} n+1 & \text{with probability} \quad (n\beta + \kappa)\,dt, \\ n & \text{with probability} \quad 1 - (n\mu + n\beta + \kappa)\,dt, \\ n-1 & \text{with probability} \quad n\mu\,dt. \end{cases}$$

---

* It appears that this sampling property of the logarithmic series distribution (which is easily proved with the aid of the generating function) has already been noticed by C. B. Williams (1947) and M. H. Quenouille.

(ii) (This is actually included in (i), but an explicit statement is desirable.) If $n = 0$, the only possible transitions in time $dt$ are from 0 to 0 or 1, the transition probabilities being

$$0 \rightarrow \begin{cases} 1 \text{ with probability } \kappa dt, \\ 0 \text{ with probability } 1 - \kappa dt. \end{cases}$$

Let $P_n(t)$ be the probability that at time $t$ the population size is $n$; then it is possible to set up an infinite system of differential-difference equations which together with the distribution $\{P_n(t_0)\}$ at some initial time $t = t_0$ determine the $\{P_n(t)\}$ at all subsequent times, and so govern the mode of growth of the population. Two alternative sets of initial conditions will be considered here:

(A) $\qquad\qquad\qquad P_0(-T) = 1 \quad \text{and} \quad P_n(-T) = 0 \quad (n > 0).$

This implies that at time $t = -T$ the population size was zero.

(B) $\qquad\qquad P_0(-T) = 0, \quad P_1(-T) = 1 \quad \text{and} \quad P_n(-T) = 0 \quad (n > 1).$

This implies that the population commenced with one individual at time $t = -T$.

Next it is necessary to give a biological interpretation of the effects associated with the constants $\beta$, $\mu$ and $\kappa$.

The first of these, $\beta$, represents the reproductive power of the individuals composing the population, the effects of sex and age being ignored. Thus it is supposed that if attention is focused on any one individual at time $t$, it will be found to undergo binary fission at a time $t + \tau$, where $\tau$ has the probability distribution

$$e^{-\beta\tau}\beta d\tau \quad (0 < \tau < \infty). \tag{8}$$

An important consequence of the assumption (8) is that the time to the next subdivision, for any individual, is statistically independent of its past history, and in particular it is independent of the length of time since that individual was itself formed by the fission of its parent. At first sight it might appear that a bacterial colony would provide a good example of such a population growing by binary fission, but it must be remembered that the generation times of bacteria, while liable to considerable random variation, have a frequency distribution* very different from (8) and possessing a pronounced non-zero mode.

The $n$ individuals present at any time are assumed to reproduce themselves independently of one another, and at the same constant mean rate. At each subdivision the parent can be thought of either as being replaced by its two offspring, or as only adding one new member to the colony and remaining a member itself; a transition $n \rightarrow n + 1$ then takes place.

In a similar way the constant $\mu$ represents the loss to the colony due to 'mortality'. It is assumed that an individual does not lose its power to reproduce unless it 'dies', and that it then ceases to be regarded as a member of the colony, so that a transition $n \rightarrow n - 1$ takes place. Such a transition could, however, also mean the removal (by any means) of an individual from the region considered; these two sources of loss are mathematically indistinguishable and will therefore be covered by the same symbol $\mu$. Thus if an individual is observed at time $t$, it will disappear from the population at a time $t + \tau$, where $\tau$ has the distribution

$$e^{-\mu\tau}\mu d\tau \quad (0 < \tau < \infty).$$

* See, for example, Kelly & Rahn (1932) and Hinshelwood (1946).

The $\beta$- and $\mu$-effects, described separately, are to be thought of as acting simultaneously and independently one of the other. Thus, when the $\beta$- and $\mu$-effects are acting together, the chance that an individual remains inactive for a time $\tau$ and then subdivides during the subsequent time interval $d\tau$ is

$$e^{-(\beta+\mu)\tau}\beta\,d\tau.$$

Integration from $\tau = 0$ to $\tau = \infty$ then gives the chance that the individual will subdivide before the $\mu$-effect has removed it from the colony; this is

$$\int_0^\infty e^{-(\beta+\mu)\tau}\beta\,d\tau = \frac{\beta}{\beta+\mu}.$$

Finally the $\kappa$-effect is one of 'immigration from outside', i.e. it is supposed that from time to time individuals not initially members of the colony may join it and proceed to behave exactly like the other members. If, from a given time instant $t$, the next such 'immigration' occurs at time $t+\tau$, the distribution of $\tau$ is assumed to be

$$e^{-\kappa\tau}\kappa\,d\tau \quad (0<\tau<\infty).$$

The structure of the model will now be clear. It only remains to point out that the probabilities of a positive unit increment from the $\beta$-effect or a negative unit increment from the $\mu$-effect in an element of time $dt$ will each be proportional to $n$, the existing population size, while the chance of a unit positive increment from the $\kappa$-effect is the same for all values of $n$.

3. The differential-difference equations of the process can now be written down. They are

$$\frac{d}{dt}P_n(t) = (n+1)\mu\,P_{n+1}(t) - \{n(\beta+\mu)+\kappa\}\,P_n(t) + \{(n-1)\beta+\kappa\}\,P_{n-1}(t), \tag{9}$$

if $n \geqslant 1$, and

$$\frac{d}{dt}P_0(t) = \mu P_1(t) - \kappa P_0(t). \tag{10}$$

It is convenient to *define* $P_n(t)$ as being identically equal to zero when $n < 0$; equation (10) can then be included within the general form (9). I owe to Dr M. S. Bartlett the remark that systems of equations of this type can most conveniently be solved with the aid of the generating function

$$\phi(z,t) \equiv \sum_{n=-\infty}^{\infty} z^n P_n(t). \tag{11}$$

It will be seen from (9) that $\phi(z,t)$ must satisfy the partial differential equation

$$\frac{\partial\phi}{\partial t} = \{\mu - (\beta+\mu)z + \beta z^2\}\frac{\partial\phi}{\partial z} + \kappa(z-1)\phi, \tag{12}$$

which together with one of the boundary conditions,

(A) $$\phi(z, -T) = 1,$$

or

(B) $$\phi(z, -T) = z,$$

and the requirement that the expansion of $\phi$ must contain no terms in $1/z, 1/z^2, \ldots$, is sufficient completely to determine the process.

The differential equation (12) is of the standard Lagrangian form, the auxiliary equations being

$$-dt = \frac{dz}{(\beta z-\mu)(z-1)} = -\frac{d\phi}{\kappa(z-1)\phi}. \tag{13}$$

'First integrals' are     $(\beta - \mu) t + \ln(z - 1) - \ln(\beta z - \mu) = \text{constant,}$

and                       $\kappa \ln(\beta z - \mu) + \beta \ln \phi = \text{constant,}$

if $\kappa > 0$ and $\beta \neq \mu$, and so the general integral of (12) is then

$$\phi(z, t) = (\mu - \beta z)^{-\kappa/\beta} \Phi \left\{ \frac{\mu - \beta z}{1 - z} e^{-(\beta - \mu)t} \right\}, \tag{14}$$

where $\Phi$ is an arbitrary function to be determined from the boundary conditions.

With boundary condition (A) it will be found that

$$\phi(z, 0) = \left( \frac{\beta - \mu}{\beta \Lambda - \mu} \right)^{\kappa/\beta} \left\{ 1 - \frac{\beta(\Lambda - 1)}{\beta \Lambda - \mu} z \right\}^{-\kappa/\beta}, \tag{15}$$

where $\Lambda$ has been written for $e^{(\beta - \mu)T}$ (this is equal to the *expected* factor by which the population will be multiplied in a time interval $T$, when it is growing in the absence of immigration). Similarly with boundary condition (B) one obtains

$$\phi(z, 0) = \frac{(\beta - \mu)^{\kappa/\beta}}{(\beta \Lambda - \mu)^{1 + \kappa/\beta}} \{ \mu(\Lambda - 1) - (\mu \Lambda - \beta) z \} \left\{ 1 - \frac{\beta(\Lambda - 1)}{\beta \Lambda - \mu} z \right\}^{-1 - \kappa/\beta}, \tag{16}$$

where $\Lambda$ has the same meaning as before.

When $\kappa = 0$, the solutions are of a slightly different form. The general integral of (12) is then

$$\phi(z, t) = \Phi \left\{ \frac{\mu - \beta z}{1 - z} e^{-(\beta - \mu)t} \right\}. \tag{17}$$

Condition (A), as might be expected, gives $\phi(z, t) \equiv 1$; this merely asserts that a zero population will remain zero if there is no immigration. Condition (B), however, gives

$$\phi(z, 0) = \frac{\mu(\Lambda - 1) - (\mu \Lambda - \beta) z}{\beta \Lambda - \mu} \left\{ 1 - \frac{\beta(\Lambda - 1)}{\beta \Lambda - \mu} z \right\}^{-1} \tag{18}$$

It will be noticed that in every case the solution is a regular function of $z$ near $z = 0$, so that in the Laurent expansion the coefficients of the negative powers will all vanish, as required.

4. It is now a simple matter to interpret these solutions. There are three cases of special interest.

(i) Consider first a population growing from zero, so that (A) is the appropriate boundary condition, the population being established in the first instance by immigration from outside. The $\kappa$-effect is of course acting all the time, even after the colony has started growing, so that in general there will at any time be present a number of independent families, each descended from a different immigrant ancestor. According to (15) the population size $n$, after the process has been developing for a time $T$, will be distributed as a negative binomial variate with index $\kappa/\beta$ and mean value

$$E(n) = \frac{\kappa(\Lambda - 1)}{\beta - \mu}. \tag{19}$$

Thus if $\beta > \mu$, the expected size of the population will for large $T$ grow geometrically at an exponential rate $(\beta - \mu)$.

If, however, $\beta < \mu$, so that the force of mortality more than compensates for the force of reproduction, one will have

$$\lim_{T \to \infty} E(n) = \frac{\kappa}{\mu - \beta}. \tag{20}$$

This is the mean of the stable distribution of population size which can just be maintained by the immigration rate $\kappa$. If $\kappa$ were equal to zero the population would almost certainly die out in a finite time.

(ii) Now suppose that $\kappa$ *is very small*, though still just greater than zero; to be more precise suppose that the ratio $\kappa/\beta$ is negligible, so that while immigration is sufficient to start off the process, and to restart it if ever the population is wiped out by an excess of the $\mu$-effect, it is negligible when compared with the contributions from the $\beta$-effect while the colony is actually growing. Then, exactly as in Fisher's analysis referred to in § 1, it will be seen that the size distribution of such a colony observed at any time will be a logarithmic series: in fact the distribution of $n$, *given that $n > 0$*, is*

$$\frac{x^n}{ny} \quad (n = 1, 2, 3, \ldots),$$

where as always $y = -\ln(1-x)$, and

$$x = \frac{\beta(\Lambda - 1)}{\beta\Lambda - \mu}. \tag{21}$$

If $T$ is very large, then

$$x \simeq 1 - \left(1 - \frac{\mu}{\beta}\right)\frac{1}{\Lambda} \quad \text{when} \quad \beta > \mu, \tag{22}$$

and

$$x \simeq \beta/\mu \quad \text{when} \quad \beta < \mu. \tag{23}$$

The second case is the more interesting, for one can then let $T$ tend to infinity and so obtain the stable distribution of population size when $\beta < \mu$. Of course in the limit when $\kappa = 0$ it is 'almost certain' that $n = 0$. If, however, a colony *does* exist (i.e. if $n > 0$), then it is almost certainly homogeneous (descended from a single immigrant ancestor), and its size $n$ will be distributed in a logarithmic series.

(iii) The case $\beta = \mu$, when reproduction and mortality just balance, requires a separate discussion. The equations auxiliary to (12) are now

$$-dt = \frac{dz}{\beta(z-1)^2} = -\frac{d\phi}{\kappa(z-1)\phi},$$

and the general integral is

$$\phi(z, t) = (1-z)^{-\kappa/\beta}\Phi\left\{\beta t + \frac{1}{1-z}\right\}.$$

Condition (A) then gives

$$\phi(z, 0) = (1 + \beta T)^{-\kappa/\beta}\left\{1 - \frac{\beta T}{1 + \beta T}z\right\}^{-\kappa/\beta}, \tag{24}$$

so that here again $n$ is distributed as a negative binomial variate with index $\kappa/\beta$, but the mean value is now

$$E(n) = \kappa T, \tag{25}$$

so that the expected size of the population is linearly proportional to the time of exposure to immigration. If $\kappa/\beta$ is very small the limiting conditional distribution of $n$ (given that $n > 0$) is once again the Fisher series

$$\frac{x^n}{ny} \quad (n = 1, 2, 3, \ldots),$$

where now

$$x = \frac{\beta T}{1 + \beta T}. \tag{26}$$

* This conditional distribution is of course obtained by taking the ratio of the general term to the sum of all the terms but the first, in the negative binomial series, and then letting $\kappa/\beta \to 0$.

5. It will perhaps have been noticed that in the analysis of the last section the immigration rate $\kappa$ (when $\kappa/\beta$ is small) merely has the effect of ensuring that an observed population is almost certainly descended from a single immigrant ancestor who entered the region considered at some instant during the time interval of length $T$ preceding the moment of observation. I think this gives us a clue to the true 'explanation' of the occurrence of the logarithmic series in the solution to the problem just considered. Before exploring the matter further it will be found helpful to examine the growth of a single family by setting $\kappa = 0$ and starting from unit population at time $t = -T$, thus employing boundary condition (B).

From the generating function (18) it then follows that the distribution of the population size $n$ at the time of observation ($t = 0$) is*

$$P_0 = \frac{\mu(\Lambda - 1)}{\beta\Lambda - \mu} \quad \text{and} \quad P_n = (1 - P_0)(1 - u)u^{n-1} \quad (n > 0), \tag{27}$$

where
$$u = \frac{\beta(\Lambda - 1)}{\beta\Lambda - \mu}.$$

The distribution is thus a geometric series with a modified zero term, the mean population size being
$$E(n) = \Lambda,$$
so that the *expected* population grows geometrically at an exponential rate $\beta - \mu$ (which will, of course, be negative if $\beta < \mu$).

It is of interest to evaluate the variance of the population size. This proves to be

$$\text{Var}(n) = E(n - \Lambda)^2 = \frac{\beta + \mu}{\beta - \mu}\Lambda(\Lambda - 1), \tag{28}$$

and the coefficient of variation of the population size is thus

$$\sqrt{\left/\left(\frac{\beta + \mu}{\beta - \mu}\right)\sqrt{}\left(1 - \frac{1}{\Lambda}\right)\right.} \quad \text{if} \quad \beta > \mu$$

and
$$\sqrt{\left/\left(\frac{\mu + \beta}{\mu - \beta}\right)\sqrt{}\left(\frac{1}{\Lambda} - 1\right)\right.} \quad \text{if} \quad \beta < \mu.$$

For large $T$ these expressions become

$$\text{C. of V.}(n) \simeq \sqrt{\left(\frac{\beta + \mu}{\beta - \mu}\right)} \quad \text{if} \quad \beta > \mu, \tag{29}$$

and
$$\text{C. of V.}(n) \simeq \sqrt{\left\{\frac{\mu + \beta}{\mu - \beta}\frac{1}{E(n)}\right\}} \quad \text{if} \quad \beta < \mu; \tag{30}$$

in the second case it will be recalled that $E(n) \to 0$ as $T \to \infty$.

The distribution of the population size thus behaves rather differently in the two cases of an exponentially growing and an exponentially decreasing population; this agrees with a conclusion reached by M. S. Bartlett (1937). He considered the similar problem when there is no spreading of generations and the generation-time is rigorously constant. As he points out, there is a connexion with Fisher's theory of the extinction of rare characters and, one might add, with the work of Francis Galton, H. W. Watson, and A. J. Lotka on the extinction of surnames.†

* (*Added in proof.*) Dr Bartlett has pointed out to me that the result (27) is stated by N. Arley & V. Borchsenius in *Acta Math.* (1945), **76**, 298–9. It is attributed by them to Dr C. Palm.

† See Fisher (1930), Galton (1889) and Lotka (1931). The problem of the distribution of surnames, and its variation in time, seems not yet to have received all the attention it deserves. The surname is a 'rare character' whose extinction can very readily be observed; normal social conditions ensure that it is inherited as if it were controlled by a gene totally sex-linked in $Y$. Reference may be made to the work of R. A. Fisher & Janet Vaughan (1939), and J. A. Fraser Roberts (1941–2), who have considered the relation between surnames and blood-groups. [See also reference to Tocher & Gray, p. 1 above. ED.]

When $\beta = \mu$, so that the forces of reproduction and mortality just balance, the above solution must be modified. The appropriate results are most easily obtained by letting $\beta$ tend to $\mu$ in the several formulae. In this way one finds

$$P_0 = \frac{\beta T}{1 + \beta T} \quad \text{and} \quad P_n = (1 - P_0)(1 - u)\, u^{n-1} \quad (n > 0), \tag{31}$$

where

$$u = \frac{\beta T}{1 + \beta T},$$

and so

$$E(n) = 1, \quad \text{Var}(n) = 2\beta T \quad \text{and} \quad \text{C. of V.}(n) = \sqrt{(2\beta T)}. \tag{32}$$

It is of interest to note that if $\beta$ and $\mu$ instead of being constants are each proportional to the *same* function of the time,* the above theory still holds, provided that $T$ is everywhere replaced by

$$\int_{-T}^{0} \psi(t)\, dt,$$

where $\beta = \beta_0 \psi(t)$ and $\mu = \mu_0 \psi(t)$.

6. Consider now the size distribution of a colony developing in the absence of immigration and known to have originated from a single individual whose arrival in the region concerned occurred during the preceding $T$ time-units. If the time of arrival of the common ancestor is a random variable uniformly distributed from $t = -T$ to $t = 0$, it will follow from (27) that

$$P_n = \frac{1}{\beta T} \int_0^T \left\{ \frac{\beta(\lambda - 1)}{\beta\lambda - \mu} \right\}^{n-1} \frac{\beta\lambda(\beta - \mu)^2}{(\beta\lambda - \mu)^2}\, d\tau \quad (n \geqslant 1),$$

where $\lambda = e^{(\beta - \mu)\tau}$. Now since

$$\frac{d}{d\tau}\left\{ \frac{\beta(\lambda - 1)}{\beta\lambda - \mu} \right\} = \frac{\beta\lambda(\beta - \mu)^2}{(\beta\lambda - \mu)^2},\; \cdot$$

this can be written

$$P_n = \frac{1}{\beta T} \int_0^U u^{n-1}\, du = \frac{U^n}{n\beta T},$$

while

$$P_0 = 1 + \frac{1}{\beta T} \ln(1 - U),$$

$U$ being the value of

$$u = \frac{\beta(\lambda - 1)}{\beta\lambda - \mu}$$

when $\tau = T$. Thus the distribution of the population size at the time of observation ($t = 0$) is

$$P_0 = 1 - \frac{1}{\beta T} \ln\left( \frac{1}{1 - x} \right) \quad \text{and} \quad P_n = (1 - P_0)\frac{x^n}{ny} \quad (n > 0), \tag{33}$$

where

$$x = \frac{\beta(\Lambda - 1)}{\beta\Lambda - \mu}, \tag{34}$$

$y$ and $\Lambda$ being defined as before. For large $T$,

$$x \simeq 1 - \left(1 - \frac{\mu}{\beta}\right)\frac{1}{\Lambda} \quad \text{if} \quad \beta > \mu, \tag{35}$$

and

$$x \simeq \beta/\mu \quad \text{if} \quad \beta < \mu. \tag{36}$$

In the latter case it is permissible to let $T$ tend to infinity and so obtain the stable (logarithmic series) distribution for $n$ (*given that $n > 0$*). When $\beta = \mu$ the $(x, T)$ relation is $x = \beta T/(1 + \beta T)$.

It will now be clear why, in § 4 (ii), the boundary condition (A) together with the hypothesis $n > 0$ led to a logarithmic series distribution for $n$ as $\kappa/\beta \to 0$. For in these circumstances it

* A discussion of the similar problem when $\beta$ and $\mu$ are *any* (not necessarily the same) functions of the time will be given in my paper (1948).

would be almost certain that an observed population was wholly descended from a single immigrant ancestor who arrived in the field of observation at an unknown time instant uniformly distributed between $t = -T$ and $t = 0$.*

7. The classical theory of population growth (see, for example, A. J. Lotka (1945) for a general review and extensive references) is largely based on a *deterministic* description of the phenomena, which leads to differential and integral equations for the expectation values of the random variables concerned (the total population size, and the numbers in the several age groups). Apart from the work on 'extinction' already mentioned, the first *stochastic* treatment of the general problems of population growth seems to be that of W. Feller in an important paper† which is unfortunately not generally accessible in this country. Feller's work has been further developed by N. Arley (1943),‡ who showed that discontinuous Markoff processes of a similar type are equally relevant in the theory of the 'cascade showers' initiated by cosmic ray particles. In particular he makes use of the simple birth-and-death process discussed here in § 5, and quotes Feller's formulae for the mean and variance of $n$ as functions of the time $t$. Arley gives an elegant method for determining an expansion for $P_n(t)$ in powers of $(t - t_0)$, and observes that the calculation of high order coefficients becomes very cumbersome. One can now see, on examining the complete solution contained in (27), that this is very reasonable, for $P_n(t)$ is quite a complicated function of the time $t$, although a simple one of the reduced variable $u$.

The results of the present paper would not have been obtained without the generating-function technique for transforming the differential-difference equations into a partial differential equation of simple type. This device was suggested to me by Dr M. S. Bartlett in the summer of 1946, and he has also applied his method to a number of Markoff processes of interest in biology, one of which is the birth-and-death process of § 5. An account of this work is now available (Bartlett, 1947). In another place I intend to give a discussion of the most general birth-and-death process of this type, in which the birth- and death-rates $\beta$ and $\mu$ can be any functions of the time $t$ (Kendall, 1948); this development also makes use of the generating-function procedure.

It is a pleasure to acknowledge my indebtedness to Dr Bartlett and to the many other friends whose comments have helped to clarify my ideas on this subject. In particular, I should like to thank Mr D. J. Finney and members of his Seminar in Biological Statistics in Oxford, and Mr F. J. Anscombe, Dr P. Jones and Dr C. B. Williams of the Rothamsted Experimental Station.

* The condition $n > 0$ is to be introduced *after* (33) has been established. If it were imposed from the start, the *a posteriori* distribution of the time of arrival of the immigrant ancestor would no longer be a uniform one, but the appropriate modification of the argument would lead to the same final result.

† Feller (1939). I am greatly indebted to Prof. Feller for the gift of a reprint of this paper. In private correspondence he tells me that he also, in unpublished lectures, has solved the equations governing the birth-and-death process discussed here in § 5.

‡ Arley's monograph is also of great value in presenting a useful account (including several new developments) of the Kolmogoroff-Feller theory of Markoff processes, especially those of the discontinuous type used here. Another excellent account has been given recently by O. Lundberg (1940). For a general introduction to this subject, reference may be made to my review article (1947).

## REFERENCES

ARLEY, N. (1943). *On the Theory of Stochastic Processes and their Application to the Theory of Cosmic Radiation.* Copenhagen: G. E. C. Gads.

BARTLETT, M. S. (1937). Deviations from expected frequencies in the theory of inbreeding. *J. Genet.* **35**, 83–7.

BARTLETT, M. S. (1947). *Stochastic Processes.* (Notes of a course given at the University of North Carolina in the Fall Quarter, 1946. It is understood that copies of these notes are available on request.)

FELLER, W. (1939). Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in Wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheoretica,* **5,** 11–40. This seems not to be generally available in Great Britain, but there is an abstract in *Math. Rev.* **1, 22.**

FISHER, R. A. (1930). *The Genetical Theory of Natural Selection,* pp. 73–83. Oxford: Clarendon Press.

FISHER, R. A. & VAUGHAN, JANET (1939). Surnames and blood-groups. *Nature, Lond.,* **144,** 1047.

FISHER, R. A., CORBET, A. S. & WILLIAMS, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12,** 42–58.

GALTON, FRANCIS (1889). *Natural Inheritance,* Appendix F (which includes the contribution by H. W. WATSON). London: Macmillan.

GREENWOOD, M. & YULE, G. U. (1920). *J. R. Statist. Soc.* **83,** 255–79.

HINSHELWOOD, C. N. (1946). *The Chemical Kinetics of the Bacterial Cell,* Chapter x (see especially Fig. 64). Oxford: University Press.

KELLY, C. D. & RAHN, O. (1932). The growth rate of individual bacterial cells. *J. Bacteriol.* **23,** 147–53.

KENDALL, D. G. (1947). A review of some recent work on discontinuous Markoff processes with applications to biology, physics, and actuarial science. *J.R. Statist. Soc.* **110,** 130–7.

KENDALL, D. G. (1948). On the generalized 'birth-and-death' process. *Ann. Math. Statist.* (in the Press).

LOTKA, A. J. (1931). The extinction of families. *J. Wash. Acad. Sci.* **21,** 377–80 and 453–9.

LOTKA, A. J. (1945). Population analysis as a chapter in the mathematical theory of evolution. Included in *Essays on Growth and Form presented to D'Arcy Wentworth Thompson* (edited by Le Gros Clark, W. E. & Medawar, P. B.). Oxford: Clarendon Press.

LUNDBERG, O. (1940). *On Random Processes and their Application to Sickness and Accident Statistics.* Uppsala: Almqvist and Wiksells.

ROBERTS, J. A. FRASER (1941–2). Blood-group frequencies in north Wales. *Ann. Eugen., Lond.,* **11,** 260–71. (Includes references to earlier work on the distribution of surnames.)

WILLIAMS, C. B. (1944*a*). Some applications of the logarithmic series and the index of diversity to ecological problems. *J. Ecol.* **32,** 1–44.

WILLIAMS, C. B. (1944*b*). The numbers of publications written by biologists. *Ann. Eugen., Lond.,* **12,** 143–6.

WILLIAMS, C. B. (1947). The logarithmic series and its application to biological problems. *J. Ecol.* (in the Press).