THE MULTIPLICATIVE PROCESS

By Richard Otter[1, 2]

University of Notre Dame

1. Introduction and summary. The multiplicative process is usually defined by the sequence of random variables $X_0, X_1, \cdots$ whose distributions are specified as follows: $P(X_0 = 1) = 1$, $\sum_{\nu=0}^{\infty} P(X_1 = \nu) = 1$, and if $X_n = 0$ then $P(X_{n+1} = 0) = 1$, whereas if X_n is a positive integer then X_{n+1} is distributed as the sum of X_n independent random variables each with the distribution of X_1. The variable X_n is interpreted as the number of "particles" in the nth generation, and the index n as a discrete time parameter. This has been the method of approach in previous studies of the process [1, 2, 3, 4, 5]. The multiplicative process has various applications, notably in the study of population growth, the spread of epidemics or rumors, and the nuclear chain reaction. The closely related "birth and death" process was recently studied by Kendall [6].

Whenever one studies the probability theory of a particular system there seem to be definite conceptual advantages in defining explicitly the set $\mathcal{T}$ of elementary events, the additive class $\mathfrak{M}$ of subsets of $\mathcal{T}$, called events, and the probability measure P for the events of $\mathfrak{M}$. Now an elementary event of this process can be represented by a rooted tree where the original particle is represented by the root vertex and where the particles of the nth generation are represented by the vertices n segments removed from the root. The tree will be finite or infinite according to whether a finite or an infinite number of particles are involved in the elementary event. Thus, the set of trees is the natural choice for $\mathcal{T}$. The first part of this paper is devoted to a more precise description of $\mathcal{T}$, $\mathfrak{M}$ and P. We shall then see easily that $X_n(t)$, the number of vertices n segments removed from the root of $t \in \mathcal{T}$, i.e. the number of particles in the nth generation, has the distribution defined in the preceding paragraph. Since the time does not appear in our description of $\mathcal{T}$ we fetter ourselves somewhat if we interpret n as a discrete time parameter. Thus, we have already reaped some harvest from considering the process from the point of view of $\mathcal{T}$. Another advantage is that we are led in a natural way to study the distribution of other structural features of the trees, e.g. the total number of vertices, or the number of vertices with k outgoing segments.

The chief results of this paper are as follows. The recursion formula for the probability P_n that a tree have n vertices $n = 1, 2, \cdots$ is obtained as well as an asymptotic estimate of P_n valid for large n. The distributions of the number of branches at the root in a finite tree, an infinite tree, or in a tree with n vertices are obtained and the asymptotic distribution of the latter as $n \to \infty$. The

---

[1] Research under an Office of Naval Research contract.

[2] The author wishes to express his gratitude to Professor E. Artin of Princeton University for the suggestion of this problem and his encouragement towards its solution.

distribution of the fraction of vertices with k outgoing segments in the finite trees, in the trees with n vertices, and the asymptotic distribution of the latter as $n \to \infty$ are also found. Finally, an estimate is obtained for the probability that a tree be finite in case this probability is near 1, a result which was previously obtained by Kolmogoroff [7].

2. The space of trees. We shall use the notation $\{a\}$, $\{a_1, a_2, \cdots a_n\}$, $\{a_j\}_{j \in J}$, and $\{a_j \mid R\}_{j \in J}$ to denote the sets which consist of respectively the single element a, the elements $a_1, a_2, \cdots a_n$, all a_j with $j \in J$, and all a_j with the property R and $j \in J$. We denote the union of two sets A and B by $A + B$, their intersection by AB, and the cartesian product of n identical factors each of which is A by $A^{(n)}$.

Let I denote the set of positive integers. We assume given for each $n \in I$ a countable set U_n of objects $u_{i_1 i_2 \cdots i_n}$ called vertices, i.e.

$$U_n = \{u_{i_1 i_2 \cdots i_n}\}_{(i_1, i_2 \cdots i_n) \in I^{(n)}} .$$

Let u_0 be a vertex distinct from all the other vertices and let $U = \{u_0\} + \Sigma\, U_n$ be the collection of all the vertices. We shall interpret u_0 as the original parent particle and the vertex u_{152}, for example, as the second son of the fifth son of the first son of the original particle. If s is a subset of U, $s \subset U$, and if $i_1, i_2, \cdots i_{n+m}$ are such that $u_{i_1 i_2, \cdots i_n}$, $u_{i_1 i_2 \cdots i_n i_{n+1}}$, $\cdots$ $u_{i_1 i_2 \cdots i_n i_{n+1} \cdots i_{n+m}}$ each belong to s then this set of vertices is called a *path* from $u_{i_1 i_2 \cdots i_n}$ to $u_{i_1 i_2 \cdots i_{n+m}}$ in s and $m \geq 0$ is the *length* of the path or the *distance* from $u_{i_1 i_2 \cdots i_n}$ to $u_{i_1 i_2 \cdots i_{n+m}}$. If $m = 1$ we call the path a *segment*, for short.

For the sake of convenience let us agree to put $u_{i_1 i_2 \cdots i_n 0} = u_{i_1 i_2 \cdots i_n}$, $(n \geq 1)$ then we define $W(s, u)$, for $u \in s \subset U$, to be the number of segments from u in s, and we call $W(s, u)$ the *type* of the vertex u in s. If t is a subset of U, then we call t a *tree* if and only if

$$W(t, u) < \infty \quad \text{for} \quad u \in t \tag{1}$$

and

$$u_{i_1 i_2 \cdots i_n} \in t \quad \text{implies} \quad u_{i_1 i_2 \cdots i_{n-1} \nu} \in t \quad \text{for} \quad \nu = 0, 1, \cdots i_n . \tag{2}$$

Let $\mathcal{T}$ be the set of all trees. The condition (2) clearly implies that for each $t \in \mathcal{T}$ we have $u_0 \in t$ and that there is a unique path from u_0 to any other vertex of t. Hence, whenever a path exists between any two vertices of t it isunique. We call u_0 the *root* of t. If for $u \in t \in \mathcal{T}$ we have $W(t, u) = 0$ then u is called an *endpoint* of t, and the vertices of t which are not endpoints are called *inner* vertices. (It is to be noted that the objects we call trees here are rooted trees in the sense of Cayley but our trees have their vertices numbered as well. Usually one would identify the trees $\{u_0, u_1, u_2, u_{11}\}$ and $\{u_0, u_1, u_2, u_{21}\}$, but we do not wish to do so because for us it is distinctly different whether the grandson is sired by the first son or by the second son.)

For $u \in t \in \mathcal{T}$ we define the *branch* of t at u to be the set of all vertices belonging to any path from u in t. Our convention of admitting paths of length 0 implies that $u \in b(t, u)$. In fact, if $W(t, u) = 0$ then $b(t, u) = \{u\}$. If t' is a tree such that $t' \subset t$ then we call t an *extension* of t', denoted $t \geq t'$ or $t' \leq t$, if $W(t', u) > 0$ implies $W(t', u) = W(t, u)$. Thus $t \geq t'$ is equivalent to $t \supset t'$ and

$$t = t' + \sum_u b(t, u)$$

where u runs through all the endpoints of t'. The extension relation imposes a partial ordering upon $\mathcal{T}$.

The extension t of t' is interpreted as a possible future aspect of a family tree when its structure at present is given by t', all present members of the family who have progeny being regarded as sterile.

If $u = u_{i_1, i_2 \cdots i_n}$ then the mapping φ defined for the vertices of $b(t, u)$ by putting

$$\varphi(u_{i_1 i_2 \cdots i_n i_{n+1} \cdots i_{n+m}}) = u_{i_{n+1} \cdots i_{n+m}}$$

maps $b(t, u)$ one to one onto a tree $\varphi(b(t, u))$ in such a fashion that if $\{v_1, v_2\}$ is a segment from v_1 to v_2 in $b(t, u)$ then $\{\varphi(v_1), \varphi(v_2)\}$ is a segment from $\varphi(v_1)$ to $\varphi(v_2)$ in $\varphi(b(t, u))$. We call the mapping φ a *homeomorphism* and we say that $b(t, u)$ is *homeomorphic* to $\varphi(b(t, u))$.

If a tree contains a finite number of vertices then it is called a *finite* tree; otherwise it is an *infinite* tree. Let $\mathcal{F}$ denote the set of all finite trees and $\mathcal{I}$ the set of all infinite trees, and let $\mathcal{K}$ denote the set of non-negative integers. For each $k \in \mathcal{K}$ we define $Y_k(t)$ for $t \in \mathcal{F}$ to be the number of vertices of type k in t. When it is clear to which tree t we refer we shall usually abbreviate $Y_0(t)$ by m, and we agree not to use the letter m with any other connotation. For each $T \in \mathcal{F}$ let $e_1(T), e_2(T), \cdots e_m(T)$ denote its m endpoints. We then define for $T \in \mathcal{F}$ and $\kappa = (k_1, k_2, \cdots k_m) \in \mathcal{K}^{(m)}$

$$[T, \kappa] = \{t \mid t \geq T, W(t, e_i(T)) = k_i, i = 1, 2, \cdots m' \in \mathcal{T}\},$$

and we call $[T, \kappa]$ a *neighborhood*. For each $t \in [T, \kappa]$ we say $[T, \kappa]$ is a *neighborhood of* t. Then it is easy to show that $\mathcal{T}$ is a *topological space* where the neighborhoods defined above form the *defining system of neighborhoods* [8].

3. The measure theory in $\mathcal{T}$. In the following paragraphs an outline of the measure theory in $\mathcal{T}$ is given which omits proofs for the most part since they are easily constructed. The only point of difficulty arises in showing the measure function to be completely additive, but here the outline has more detail.

Let $\mathfrak{S}$ be the collection of subsets of $\mathcal{T}$ such that $0 \in \mathfrak{S}$ and any other set S belongs to $\mathfrak{S}$ if and only if there is a $t \in \mathcal{F}$ and a non-void "rectangle set" $A = A_1 \times A_2 \times \cdots A_m \subset \mathcal{K}^{(m)}$, $m = Y_0(t)$, such that

$$S = \sum_{\kappa \in A} [t, \kappa] \tag{3}$$

where the sets $A_1, A_2, \cdots A_m$ may be finite or infinite sets of non-negative integers. The collection of neighborhoods which appear as terms in (3), i.e. $\{[t, \kappa]\}_{\kappa \epsilon A}$, we call an $\mathfrak{S}$-*partition* of S, and t is called the *generator* of the $\mathfrak{S}$-partition. Only a finite number of $\mathfrak{S}$-partitions are possible for an $S \epsilon \mathfrak{S}$, because only a finite number of trees can possibly be generators and there is only one $\mathfrak{S}$-partition per generator. With respect to our partial ordering of the trees all possible generators lie between two particular ones. We call the smaller of these the *irreducible generator* and the corresponding $\mathfrak{S}$-partition the *irreducible* $\mathfrak{S}$-*partition* of S. Any partition of S into neighborhoods must be a subpartition of this irreducible $\mathfrak{S}$-partition. The elements of $\mathfrak{S}$ also display two important properties of the rectangles in Euclidean space, namely if S $S' \epsilon \mathfrak{S}$ then

$$SS' \epsilon \mathfrak{S} \tag{4}$$

and if $S \subset S'$ then there is a finite chain

$$S = S_0 \subset S_1 \subset \cdots S_n = S' \tag{5}$$

such that $S_i, S_i - S_{i-1} \epsilon \mathfrak{S}$ for $i = 1, 2, \cdots n$.
A class of sets with the properties (4) and (5) has been called a *half-ring* by von Neumann [9].

Let $p_0, p_1, \cdots$ be given non-negative numbers such that $\sum_0^\infty p_\nu = 1$. For $t \epsilon \mathcal{T}$ let us put

$$\bar{P}(t) = \prod_{\nu=1}^{\infty} p_\nu^{Y_\nu(t)} \tag{6}$$

with the convention $0^0 = 1$. We then define the measure function P for the sets in $\mathfrak{S}$ by

$$P(0) = 0$$

$$P([t, \kappa]) = \left(\prod_{\nu=1}^{m} p_{\kappa_\nu}\right)\bar{P}(t), \quad \text{where } \kappa = (k_1, k_2 \cdots k_m) \epsilon \mathcal{K}^{(m)} \tag{7}$$

$$P(S) = \sum_{\kappa \epsilon A} P([t, \kappa]), \quad \text{where } \{[t, \kappa]\}_{\kappa \epsilon A} \text{ is the irreducible } \mathfrak{S}\text{-partition of } S.$$

P is evidently non-negative. Letting t be the tree with one vertex and putting $A = \mathcal{K}$ gives $P(\mathcal{T}) = 1$. It is easy to see that P is completely additive for the $\mathfrak{S}$-partitions of a neighborhood, but this implies P is completely additive for the $\mathfrak{S}$-partitions of an arbitrary element S of $\mathfrak{S}$. In order to show that P is completely additive for any partition of S into elements of $\mathfrak{S}$, it is necessary and sufficient to show this for an arbitrary partition of a neighborhood into neighborhoods. One may reach finer and finer partitions of a given neighborhood N by replacing a neighborhood in any one partition by an $\mathfrak{S}$-partition of the neighborhood, and repeating the process. The sum of the measures of the sets in the partition is invariant under such a replacement. On the other hand it can be shown that all possible partitions of N into neighborhoods may be reached in this way. More precisely, let $\bar{N} = \{N_j\}_{j \epsilon J}$ be a partition of a neighborhood N

into neighborhoods N_j. We call $\bar{N}$ reduced if whenever a subset of $\bar{N}$ is an $\mathfrak{S}$-partition of a neighborhood $M \subset N$ then the partition consists of M itself, i.e. it is the irreducible $\mathfrak{S}$-partition of M. Then we have the following theorem:

THEOREM 1. *If $\bar{N}$ is a reduced partition of a neighborhood N into neighborhoods then $\bar{N} = \{N\}$.*

The proof is indirect and proceeds by constructing a decreasing sequence of neighborhoods contained in N whose limit is not void and yet has nothing in common with any N_j, but this is a contradiction.

Let $\mathfrak{F}$ consist of all those sets which may be formed by finite unions of disjoint elements of a half-ring $\mathfrak{S}$, then $\mathfrak{F}$ is a field of sets. If P is a completely additive measure on $\mathfrak{S}$ then its natural extension P_1 is completely additive on $\mathfrak{F}$ [9]. Kolmogoroff [10] has shown that the completely additive measure P_1 may always be extended to a completely additive measure P_2 on the Borel field $\mathfrak{M}$, i.e. the smallest additive class of sets containing $\mathfrak{F}$. Since $P_2(\mathcal{T}) = 1$, P_2 is a probability measure. For simplicity we put $P_2 = P$. Let us also agree that if M is the set of all trees with the property R we may write $P(R)$ instead of $P(M)$. If N is a set with $P(N) > 0$ then $P(M/N)$ shall denote the conditional probability of M, given N, i.e. $P(M/N) = (P(N))^{-1}P(MN)$.

4. Independence of the branches. In the multiplicative process the events occurring in one branch of a tree are independent of those in a second branch disjoint with the first and it is for this reason that the process is relatively simple to analyze. In this section we shall try to expose the character of this independence.

For $T \in \mathcal{F}$, let $\mathcal{E}_t$ be the set of all extensions of T, then

$$\mathcal{E}_T = \sum_{\kappa \in \mathcal{K}^{(m)}} [T, \kappa],$$

whence by (6) and (7) $P(\mathcal{E}_T) = \bar{P}(T)$. The following lemma is then easily established.

LEMMA 1. *If $P(\mathcal{E}_T) > 0$ then $W(t, e_i(T))$, $i = 1, 2, \cdots m$, under the condition $t \in \mathcal{E}_T$, are independent-random variables each with the distribution,*

$$(8) \qquad P(W(t, e_i(T)) = k/\mathcal{E}_T) = p_k \qquad k = 0, 1, 2, \cdots .$$

In the particular case where $T = \{u_0\}$ we have $\mathcal{E}_T = \mathcal{T}$ and we put $W(t) = W(t, u_0)$ for short. Thus $W(t)$ tells what type of vertex the root of t is and (8) becomes

$$P(W = k) = p_k \qquad k = 0, 1, 2, \cdots ,$$

For $t \in \mathcal{T}$ and $n = 0, 1, 2, \cdots$ let $X_n(t)$ be the number of vertices of t at distance n from its root. Then $X_0(t) = 1$ and $X_1(t) = W(t)$. If n, r are positive integers then there is at least one $T \in \mathcal{F}$ which has r of its endpoints, say $e_{i_1}(T), e_{i_2}(T), \cdots e_{i_r}(T)$, at distance n from the root and which also satisfies $X_{n+1}(T) = 0$. Put

$$\mathcal{E}_T^{i_1 \cdots i_r} = \{t \mid W(t, e_j(T)) = 0, j \neq i_1, i_2, \cdots i_r, t \in \mathcal{E}_T\}.$$

Evidently for $t \in \mathcal{E}_T^{i_1 \cdots i_r}$

$$X_{n+1}(t) = \sum_{\nu=1}^{i} W(t, e_{i_\nu}(T)),$$

and a proof similar to that of lemma 1 gives

LEMMA 2. *If* $P(\mathcal{E}_T^{i_1 \cdots i_r}) > 0$ *then* $X_{n+1}(t)$, *under the condition* $t \in \mathcal{E}_T^{i_1 \cdots i_r}$, *is the sum of r independent random variables each with the distribution of* X_1.

By (6) and (7) for $t \in \mathfrak{M} \subset \mathcal{F}\mathcal{E}_T$

$$P(t) = \prod_{\nu=0}^{\infty} p_\nu^{Y_\nu(t)},$$

which depends only upon the type of each vertex as it occurs in t. For those vertices which are inner vertices of T, $Y_\nu(t)$ is constant. Any other vertex belongs to one and only one of $b(t, e_1(T))$, $b(t, e_2(T))$, $\cdots$ $b(t, e_m(T))$ and its type in t is, of course, the same as its type in the branch to which it belongs. Furthermore, each branch is homeomorphic to just one tree in $\mathcal{F}$,

$$b(t, e_i(T)) \leftrightarrow t_i, \qquad i = 1, 2, \cdots m.$$

Since the type of a vertex is preserved under homeomorphism we have

$$P(t) = P(\mathcal{E}_T)P(t_1)P(t_2) \cdots P(t_m).$$

If, as t runs through $\mathfrak{M}$, $(t_1, t_2, \cdots t_m)$ runs through $\mathfrak{M}_1 \times \mathfrak{M}_2 \times \cdots \mathfrak{M}_m$, we obtain

$$P(\mathfrak{M}) = P(\mathcal{E}_T)P(\mathfrak{M}_1)P(\mathfrak{M}_2) \cdots P(\mathfrak{M}_m). \tag{10}$$

Let us hereafter put $p = P(\mathcal{F})$. In the particular case of (10) where $\mathfrak{M} = \mathcal{F}\mathcal{E}_T$ we clearly have $\mathfrak{M}_i = \mathcal{F}$, $i = 1, 2, \cdots m$, hence

$$P(\mathcal{F}\mathcal{E}_T) = P(\mathcal{E}_T)\cdot p^m. \tag{11}$$

If we define T_ν, $\nu = 0, 1, 2, \cdots$, to be the tree with $\nu + 1$ vertices which has $W(T_\nu) = \nu$ then

$$\begin{aligned} \mathcal{T} &= \{u_0\} + \sum_{\nu=1}^{\infty} \mathcal{E}_{T_\nu}, \\ \mathcal{F} &= \{u_0\} + \sum_{\nu=1}^{\infty} \mathcal{F}\mathcal{E}_{T_\nu}, \end{aligned} \tag{12}$$

where

$$\mathcal{E}_{T_i}\mathcal{E}_{T_j} = \mathcal{E}_{T_j}\{u_0\} = 0, \qquad i \neq j;$$

$$P(\mathcal{E}_{T_\nu}) = p_\nu, \qquad \nu = 1, 2, \cdots.$$

From (11) and (12) we get

$$\sum_{\nu=0}^{\infty} p_\nu p^\nu = p. \tag{13}$$

For $t \in \mathcal{F}\mathcal{E}_T$ let $Z(b(t, e_i(T)))$ be the number of vertices in the branch of t at $e_i(T)$. In the particular case where $T = \{u_0\}$ we have $b(t, u_0) = t$ and $Z(t)$ is the number of vertices of t. If now

$$\mathcal{F}_n = \{t \mid Z(t) = n, t \in \mathcal{F}\}; \qquad n = 1, 2, \cdots ;$$
$$P_n = P(\mathcal{F}_n),$$

then by putting $\mathfrak{M} = \mathfrak{M}_T^{n_1 \cdots n_m}$ where

$$\text{(14)} \quad \begin{aligned} \mathfrak{M}_T^{n_1 \cdots n_m} &= \{t \mid Z(b(t, e_i(T))) = n_i , \qquad i = 1, 2, \cdots m, t \in \mathcal{F}\mathcal{E}_T\}, \\ \mathfrak{M}_i &= \mathcal{F}_{n_i} , \qquad i = 1, 2, \cdots m, \end{aligned}$$

we may apply (10), which gives

$$\text{(15)} \quad P(t \in \mathcal{F}\mathcal{E}_T , Z(b(t, e_i(T))) = n_i , i = 1, 2, \cdots m) = P(\mathcal{E}_T)P_{n_1}P_{n_2} \cdots P_{n_m} .$$

If $p > 0$ we may multiply and divide the right hand member of (15) by p^m which leads us to the following lemma:

LEMMA 3. *If $P(\mathcal{F}\mathcal{E}_T) > 0$, then $Z(b(t, e_i(T)))$, $i = 1, 2, \cdots m$, under the condition $t \in \mathcal{F}\mathcal{E}_T$, are m independent random variables each with the distribution of $Z(t)$, given $t \in \mathcal{F}$.*

5. The distribution of $Z(t)$. Let $f(w)$ be the generating function for the distribution of W, i.e.

$$\text{(16)} \qquad f(w) = \sum_{\nu=0}^{\infty} p_\nu w^\nu$$

where w is a complex variable. If one is interested in studying the sequence $X_0 , X_1 , \cdots$ then one should define another sequence of functions $f_0 , f_1 , \cdots$ where $f_0(w) = w$ and $f_{n+1}(w) = f(f_n(w))$ for $n = 0, 1, 2, \cdots$. By computing formally the expansion of $f_n(w)$ around $w = 0$ it is not difficult to show that $f_n(w)$ is the generating function for X_n , i.e. $f_n(w) = \sum_{\nu=0}^{\infty} P(X_n = \nu)w^\nu$ which is the starting point for the previous investigations of the multiplicative process. But since we shall be mainly interested in the distribution of Z we define $\mathcal{P}(z)$ to be the corresponding generating function, i.e.

$$\text{(17)} \qquad \mathcal{P}(z) = \sum_{n=1}^{\infty} P_n z^n .$$

Let ρ and α be the radii of convergence of the power series in the right members of (16) and (17) respectively. Since $f(1) = 1$ and $\mathcal{P}(1) = P(\mathcal{F}) \leq 1$ we know $\rho, \alpha \geq 1$ hence $f(w)$ and $\mathcal{P}(z)$ are analytic in $| w | < \rho$ and $| z | < \alpha$ respectively. The relation between the distribution of W and that of Z is put in evidence by the following theorem:

THEOREM 2. *Let*

$$\mathcal{G}(z, w) = zf(w) - w,$$

then $w = \mathcal{P}(z)$ *is the unique analytic solution of*

$$\mathcal{G}(z, w) = 0 \tag{18}$$

in a certain neighborhood of (0, 0).

PROOF. Since $\mathcal{P}(z)$ is analytic at 0 and $\mathcal{P}(0) = 0$ it suffices to show that if we substitute formally $\sum P_n Z^n$ for w in $z \sum p_\nu w^\nu$ the coefficient of z^n is uniquely determined and is P_n.

$$z \sum_{\nu=0}^{\infty} p_\nu (\mathcal{P}(z))^\nu = p_0 z + \sum_{n=2}^{\infty} \left(\sum_{\nu=1}^{\infty} \sum_{\Sigma n_i = n-1} p_\nu P_{n_1} P_{n_2} \cdots P_{n_\nu} \right) z^n. \tag{19}$$

If in (14) we put $T = T_\nu$, where T_ν was defined just before (12), then $m = Y_0(T_\nu) = \nu$. Let us require in addition that the total number of vertices in the branches be $n - 1$, i.e. $n_1 + n_2 + \cdots n_\nu = n - 1$, then

$$\mathcal{F}_n = \sum_{\nu=1}^{\infty} \sum_{\Sigma n_i = n-1} \mathfrak{M}_{T_\nu}^{n_1 \cdots n_\nu}, \qquad n = 2, 3, \cdots, \tag{20}$$

where

$$\mathfrak{M}_{T_i}^{n_1 \cdots n_i} \cdot \mathfrak{M}_{T_j}^{m_1 \cdots m_j} = 0,$$

unless $i = j$ and $n_1 = m_1$, $n_2 = m_2$, $\cdots n_j = m_j$. By applying P to (20) and using (15) we get the coefficient of Z^n in (19) for $n \geq 2$. This together with the obvious fact that $P_1 = p_0$ completes the proof.

It is worthwhile noticing that by means of the formula of Burman and Lagrange [11] we can solve the recursion formula for P_n in terms of $p_0, p_1, \cdots$, namely

$$P_n = \frac{1}{n!} \left[\frac{d^{(n-1)}}{dw^{n-1}} (f(w))^n \right]_{w=0} = \sum_{\substack{\Sigma \nu_j = n \\ \Sigma j \nu_j = n-1}} \frac{(n-1)!}{\nu_0! \nu_1! \cdots} p_0^{\nu_0} p_1^{\nu_1} \cdots. \tag{21}$$

Now if t has n vertices we know from Euler's characteristic that $\sum j Y_j(t) = n - 1$. Since $P(t) = \prod p_j^{Y_j(t)}$ we see from (21) that

$$\frac{(n-1)!}{\nu_0! \nu_1! \cdots}, \qquad \Sigma \nu_j = n, \qquad \Sigma j \nu_j = n - 1,$$

is the number of trees in $\mathcal{F}_n$ for which $Y_0(t) = \nu_0$, $Y_1(t) = \nu_1, \cdots$.

Evidently $w = \mathcal{P}(z)$ remains a solution of (18) for all z such that $|z| < \alpha$, $|w| < \rho$. In case $p_0 = 0$ the constant 0 solves (18). Hence $\mathcal{P}(z) = 0$ for all z and so $\mathcal{P}(1) = p = 0$. Conversely, if $p = 0$ then $P_1 = p_0 = 0$ which gives

COROLLARY 1. $p = 0$ *if and only if* $p_0 = 0$.

Since we wish to investigate the distribution in $\mathcal{F}$ we shall henceforth assume $p_0 \neq 0$.

Any non-constant function $g(z)$ which has a power series development possessing non-negative coefficients $g(z) = \sum a_\nu z^\nu$, $a_\nu \geq 0$ with a positive radius of convergence R has two properties that are important for us:

(22) $g(z)$ has a singularity at R.

(23) If $\sum a_\nu R^\nu$ converges then $\sum a_\nu z_0^\nu$ converges absolutely and uniformly for $|z_0| = R$, and so the series defines a continuous function $g(z)$ there. We have $\lim_{z \to z_0} g(z) = \sum a_\nu z_0^\nu$ as long as the path of approach to z_0 lies in $|z| \leq R$. On the other hand, if as z approaches R through real values below R, $z \to R-$, the limit of $g(z)$ exists then $\sum a_\nu R^\nu$ converges. So if we put $g(R) = \lim_{z \to R-} g(z) = \sum a_\nu R^\nu$ then the meaning is unique even allowing ∞ as a value.

Returning to $\mathcal{P}(z)$, if for $|z| \leq \alpha$ we have $|w| \leq \rho$ where $w = \mathcal{P}(z)$, then

$$z = \frac{w}{f(w)} = \mathcal{P}^{-1}(w), \tag{24}$$

which shows the mapping is schlicht in such a domain and that the image domain cannot contain zeros of $f(w)$. Because of (23) and the fact that $\mathcal{P}(1)$ is finite even if $\alpha = 1$ we see that the mapping is certainly one to one for $|z| \leq 1$.

Corollary 2. *p is the smallest root of $f(w) = w$ in $0 \leq w \leq 1$.*

Proof. (13) shows p is a root in the interval. If for $0 \leq w_0 \leq p$ we have $f(w_0) = w_0$ then by (24) $\mathcal{P}^{-1}(w_0) = 1$.

The following corollary is the well known criterion for extinction

Corollary 3. *$p = 1$ if and only if $f'(1) \leq 1$.*

Proof. $p = 1$, $p_0 > 0$, and the convexity of $f(w)$ in $0 \leq w < 1$ guarantee that $(f(w) - 1)/(w - 1)$ is bounded by 1 and is monotonic increasing with w. Hence $f'(1)$ exists and is ≤ 1.

Conversely, if $f'(1) \leq 1$ then either $f'(w)$ is constant $(= p_1 < 1)$ in $0 \leq w < 1$ or else it is strictly increasing with w and in either case $f'(w) < 1$. The mean value theorem gives $f(w) > w$ in $0 \leq w < 1$, hence $p = 1$.

Putting $a = \mathcal{P}(\alpha)$ we have the following lemma:

Lemma 4. *$a \leq \rho$.*

Proof. We already know that $\mathcal{P}(z)$ has a unique analytic inverse given by (24) for $|\mathcal{P}(z)| < \rho$, but on the other hand $\mathcal{P}'(z) \neq 0$ for $0 \leq z < \alpha$ so this inverse is analytic for $0 \leq w < a$. If we had $a > \rho$ we could continue $f(w)$ analytically by means of (24) along the real axis past its singularity at ρ, but this is impossible.

Corollary. *$p = 1$ if and only if $a \geq 1$.*

Proof. The necessity follows from the monotone behavior of $\mathcal{P}(z)$ for $0 \leq z < \alpha$. Conversely, if $a \geq 1$ then $z = \mathcal{P}^{-1}(1) = 1$.

Theorem 3. *If $p_0 + p_1 \neq 1$, then*

(25) *α and a are finite;*

(26) $f(a) = a/\alpha$;

(27) *$f'(a) \leq 1/\alpha$ where the strict inequality can hold only if $a = \rho$.*

Proof. Let $r \geq 2$ be such that $p_r \neq 0$, then for $0 < z < \alpha$, we get from the functional equation

$$zp_r(\mathcal{P}(z))^r - \mathcal{P}(z) < 0;$$

$$0 < \mathcal{P}(z) < \left(\frac{1}{zp_r}\right)^{1/(r-1)}.$$

By letting $z \to \alpha -$ we see α is finite and $\mathcal{P}(z)$ is bounded. Since $\mathcal{P}(z)$ is monotonic in this region we get $a < \infty$. By letting $z \to \alpha$ in $\mathcal{G}(z, \mathcal{P}(z))$ we get (26). For $0 \leq z \leq \alpha$, $\mathcal{G}_w(z, \mathcal{P}(z)) = zf'(\mathcal{P}(z)) - 1$ is continuous and monotonic increasing with z and is < 0 for z near 0. From the general theorem on implicit functions we know $\mathcal{G}_w(z, \mathcal{P}(z)) \neq 0$ for $|z| < \alpha$, so if we let $z \to \alpha$ we obtain (27).

If $a = \rho$ (27) merely guarantees the finiteness of $f'(\rho)$ and gives an upper bound. One can easily construct an example where $1/\alpha$ is the least upper bound and one where it is not.

But if $a < \rho$ then since $\mathcal{G}(z, w)$ is analytic at (α, a) and $\mathcal{G}(\alpha, a) = 0$ we obtain from the implicit function theorem the strict equality in (27).

COROLLARY. *If* $\alpha = 1$ *then* $a = p = 1$.

PROOF. By (26)

$$f(a) = a = \mathcal{F}(1) = p \leq 1. \tag{28}$$

If $a < \rho$ then $f'(p) = 1$ so $p = 1$ from the convexity of $f(w)$. If $a = \rho$ then $a \geq 1$ which when combined with (28) gives $a = 1$.

The case where $p_0 + p_1 = 1$ escapes Theorem 3 but it is easily examined separately, namely

$$f(w) = p_0 + p_1 w, \qquad p_0 \neq 0,$$

$$\mathcal{P}(z) = \sum_{n=1}^{\infty} p_0 p_1^{n-1} z^n = \frac{p_0 z}{1 - p_1 z}.$$

Hence $p = 1$, $\alpha = 1/p_1$ and $a = \rho = \infty$.

For the practical applications of the theory it is valuable to know some conditions which guarantee $a < \rho$, and thus strict equality in (27). From the foregoing analysis it is evident that one such condition is $\rho = \infty$, i.e. $f(w)$ is an entire function, and another is $f'(1) > 1$. If one has enough information about $f(w)$ to plot its graph for real positive w then the line through the origin tangent to $f(w)$ in the first quadrant touches the curve at the point $(a, a/\alpha)$ from which we determine both a and α.

6. Asymptotic properties of the distributions. If we examine the terms of the sequence $p_0, p_1, \cdots$ we may find that the indices of the non-zero terms are all multiples of some common integer larger than 1. In this case we should expect to have $P_n = 0$ with the same sort of regularity. So let us define q to be the largest integer such that $p_\nu \neq 0$ implies ν is a multiple of q. Clearly we have $q \geq 1$ and $q = 1$ means there is no integer other than 1 which divides the indices of all the non-zero p_ν. Of course, $p_1 \neq 0$ implies $q = 1$. The following theorem establishes an asymptotic estimate for P_n valid for large n, provided

$n - 1$ is a multiple of q, and incidentally shows that $P_n = 0$, if $n - 1$ is not a multiple of q.

THEOREM 4. *If $a < \rho$ then*

$$(29) \qquad P_n = \begin{cases} q\left(\dfrac{a}{2\pi\alpha f''(a)}\right)^{\frac{1}{2}} \alpha^{-n} n^{-\frac{3}{2}} + O(\alpha^{-n} n^{-\frac{5}{2}}), & n \equiv 1 \pmod q; \\ 0 \qquad , & n \not\equiv 1 \pmod q, \end{cases}$$

i.e. for large $n \equiv 1 \pmod q$

$$P_n \sim q\left(\frac{a}{2\pi\alpha f''(a)}\right)^{\frac{1}{2}} \alpha^{-n} n^{-\frac{3}{2}}.$$

PROOF. Let us put $\theta = 2\pi/q$, then for $|w| \leq a$,

$$|f(w)| = \left|\sum_{k=0}^{\infty} p_{kq} w^{kq}\right| \leq \sum_{k=0}^{\infty} p_{kq} |w|^{kq} = f(|w|),$$

and the equality evidently holds if and only if arg w is an integral multiple of θ. Furthermore, if w is such that $|f(w)| = f(|w|)$ and we put $z = \mathcal{P}^{-1}(w)$ then $w = \mathcal{P}(w/f(w))$ so we get

$$|\mathcal{P}(z)| = \mathcal{P}\left(\frac{|w|}{f(|w|)}\right) = \mathcal{P}\left(\frac{|w|}{|f(w)|}\right) = \mathcal{P}(|z|),$$

hence $P_n = 0$, if $n \not\equiv 1 \pmod q$.

For $|z| = \alpha$ and $w = \mathcal{P}(z)$ the point (z, w) satisfies (18) by (23). If we put

$$z_\nu = \alpha e^{i\nu\theta},$$
$$w_\nu = a e^{i\nu\theta}, \qquad \nu = 0, 1, \cdots q - 1,$$

then $w_\nu = \mathcal{P}(z_\nu)$ and

$$\mathcal{G}_w(z_\nu, w_\nu) = z_\nu f'(w_\nu) - 1 = \alpha f'(a) - 1 = 0,$$

so that $z_0, z_1, \cdots z_{q-1}$ are certainly singularities of $\mathcal{P}(z)$. But $f(w)$ is analytic at w_ν and $f(w_\nu) = a/\alpha \neq 0$, so the solution of (18) for z,

$$z = \mathcal{P}^{-1}(w) = \frac{w}{f(w)},$$

is analytic at w_ν. Furthermore

$$\frac{d}{dw}\mathcal{P}^{-1}(w_\nu) = \frac{1 - z_\nu f'(w_\nu)}{f(w_\nu)} = 0,$$

$$\frac{d^2}{dw^2}\mathcal{P}^{-1}(w_\nu) = -\frac{z_\nu f''(w_\nu)}{f(w_\nu)} = -\frac{\alpha^2 f''(a)}{w_\nu} \neq 0,$$

which shows that $\mathcal{P}(z)$ has a branch point of order 1 at each z_ν, i.e. $\mathcal{P}(z)$ is an analytic function of $(z - z_\nu)^{1/2}$ in the neighborhood of (z_ν, w_ν), $\nu = 0, 1, \cdots q - 1$.

For $|z| = \alpha$, $w = \mathcal{P}(z)$ but $z \neq z_\nu$ we obtain

$$|\mathcal{G}_w(z, w)| > 1 - \alpha|f'(w)| > 1 - \alpha f'(|w|) > 1 - \alpha f'(a) = 0,$$

hence $\mathcal{P}(z)$ is an analytic function of z in a certain neighborhood of such a pair (z, w).

By analytic continuation we find a circle of radius $\beta > \alpha$ such that $\mathcal{P}(z)$ is an analytic function of $(z - z_\nu)^{1/2}$ for $|z| \leq \beta$. If we make radial cuts in this circle running outward from each z_ν then in the resulting domain D each of the functions $(z - z_\nu)^{1/2}$ is an analytic function of z hence so is $\mathcal{P}(z)$.

Let Γ be the path consisting of the boundary of D oriented in the positive sense, let γ be the part of Γ lying in the sector $-\pi/q \leq \arg z \leq \pi/q$, and let γ' be that part of γ leading from β to α along the lower lip of the cut at α, thence along the upper lip back to β. Since $\mathcal{P}(z)$ satisfies the relation $\mathcal{P}(e^{i\nu\theta}z) = e^{i\nu\theta}\mathcal{P}(z)$ for $\nu = 0, 1, \cdots q - 1$, we see from Cauchy's formula that

$$P_n = \frac{1}{2\pi i}\int_\Gamma \frac{\mathcal{P}(z)}{z^{n+1}}\,dz = \frac{A}{2\pi i}\int_\gamma \frac{\mathcal{P}(z)}{z^{n+1}}\,dz,$$

where

$$\begin{aligned} A = \sum_{\nu=0}^{q-1} e^{-i\nu\theta(n-1)} &= 0, \qquad n \not\equiv 1 \pmod{q}; \\ &= q, \qquad n \equiv 1 \pmod{q}. \end{aligned}$$

Restricting ourselves to $n \equiv 1 \pmod{q}$ we put

$$\mathcal{P}(z) = a + b(z - \alpha)^{1/2} + c(z - \alpha) + (z - \alpha)^{3/2}\mathcal{Q}(z),$$

where $\mathcal{Q}(z)$ is analytic in D. Then $P_n = B + C$, where

$$B = \frac{q}{2\pi i}\int_\gamma \frac{a + b(z - \alpha)^{\frac{1}{2}} + c(z - \alpha)}{z^{n+1}}\,dz,$$

$$C = \frac{q}{2\pi i}\int_\gamma \frac{(z - \alpha)^{\frac{3}{2}}\mathcal{Q}(z)}{z^{n+1}}\,dz.$$

We find

$$B = \frac{bq}{2\pi i}\int_\Gamma \frac{(z - \alpha)^{\frac{1}{2}}}{z^{n+1}}\,dz + O(\beta^{-n}) = ibq\sqrt{\alpha}(-1)^n\binom{1/2}{n}\alpha^{-n} + O(\beta^{-n});$$

$$|C| = O\left(\int_{\gamma'} \frac{|z - \alpha|^{\frac{3}{2}}}{|z|^{n+1}}\,|dz|\right) = O\left(\left|\int_{\gamma'} \frac{(z - \alpha)^{\frac{3}{2}}}{z^{n+1}}\,dz\right|\right) = O\left(\alpha^{-n}\left|\binom{3/2}{n}\right|\right).$$

The constant b is determined from the equations

$$w - a = b(z - \alpha)^{1/2} + \cdots;$$

$$z - \alpha = -\frac{\alpha^2 f''(a)}{2a}(w - a)^2 + \cdots.$$

Using the fact that

$$\left|\binom{1/2}{n}\right| = (4\pi n^3)^{-1/2} + O(n^{-5/2}),$$

$$\left|\binom{3/2}{n}\right| = O(n^{-5/2}),$$

we finally obtain (29) as desired.

Thus P_n approaches zero a little faster than exponentially with n regardless of whether $p = 1$ or $p < 1$, except for the special case when $\alpha = 1$. In this case it is interesting that, according to the corollary to lemma 4, $p = 1$.

The case where $q \neq 1$ is of no practical importance since one can always bring q back to 1 by making a very small decrease in one of the non-zero p_ν and increasing p_1 by the same amount. This can clearly be done so that none of the important characteristics of $f(w)$ is changed appreciably.

7. The limiting distributions of $W(t)$ and $n^{-1}Y_k(t)$ for $t \in \mathcal{F}_n$. Let us momentarily drop the condition $p_0 \neq 0$. The characteristic function of W is

$$\int_{\mathfrak{I}} e^{i\theta W}\, dP = f(e^{i\theta}), \tag{30}$$

so that for the rth moment of W we have

$$E(W^r) = \frac{d^{(r)}}{d(i\theta)^r} f(e^{i\theta})\Big|_{\theta=0}, \qquad r = 0, 1, 2, \cdots. \tag{31}$$

For the first and second moments we obtain

$$E(W) = f'(1),$$
$$E(W^2) = f'(1) + f''(1),$$

which shows that the criterion for extinction (Corollary 3 to Theorem 2) may be stated as follows: the multiplicative process is almost certain to expire if and only if $E(W) \leq 1$. From (30) we see that all the moments of W will be finite as soon as $\rho > 1$; but if $\rho = 1$ no general statement can be made, except in case $a = 1$ also, for indeed $a = 1$ implies $\alpha = 1$ so by (31) and (27) $E(W) = f'(1) \leq 1$.

We now reassume $p_0 \neq 0$. Since the variables $Z, Y_0, Y_1, \cdots$ are restricted to $t \in \mathcal{F}$ it is convenient to see what happens to W in $\mathcal{F}$. If we define $g(w) = p^{-1}f(pw)$ then (13) shows $g(w)$ and $g(e^{i\theta})$ are the generating function and characteristic function respectively for W, given $t \in \mathcal{F}$. Thus we see immediately that the first moment of W, given $\mathcal{F}$, is always ≤ 1, and all its moments are finite if $p < 1$.

In case $0 < p < 1$ we may also introduce $h(w)$ defined by

$$f(w) = pg(w) + (1 - p)h(w), \tag{32}$$

then $h(w)$ is obviously the generating function of W, given $\mathcal{G}$. Here the rth moment is finite whenever the rth moment of W is finite. (32) gives

$$P(W = k/\mathcal{G}) = p_k \frac{1 - p^k}{1 - p}, \qquad k = 1, 2, \cdots.$$

It would be interesting to be able to compare this with the corresponding thing for large finite trees and in this connection we have the following theorem:

THEOREM 5. *If $a < \rho$ and $q = 1$,*

$$\lim_{n \to \infty} P(W = k/\mathcal{F}_n) = \alpha k p_k a^{k-1}, \qquad k = 1, 2, \cdots.$$

PROOF. By expanding $zf(e^{i\theta}\mathcal{P}(z))$ in powers of z we obtain

$$zf(e^{i\theta}\mathcal{P}(z)) = \sum_{n=1}^{\infty} \phi_n(\theta)z^n, \tag{33}$$

where

$$\phi_n(\theta) = \int_{F_n} e^{i\theta W} dP = \sum_{\nu=1}^{\infty} \sum_{\Sigma n_j = n-1} e^{i\nu\theta} p_\nu P_{n_1} P_{n_2} \cdots P_{n_\nu},$$

so that if $P_n \neq 0$ then $P_n^{-1}\phi_n(\theta)$ is the characteristic function of W, given $\mathcal{F}_n$. From (33) we get

$$\phi_n(\theta) = \frac{1}{2\pi i} \int_\Gamma \frac{f(e^{i\theta}\mathcal{P}(z))}{z^n} dz.$$

Since $a < \rho$ we may expand $f(e^{i\theta}\mathcal{P}(z))$ about the point $\mathcal{P}(z) = a$ and integrate as in the proof of theorem 4, thus

$$P_n^{-1}\phi_n(\theta) = \frac{P_{n-1}}{P_n} e^{i\theta} f'(ae^{i\theta}) + \epsilon_n(\theta).$$

Since $\epsilon_n(\theta) \to 0$ as $n \to \infty$,

$$\lim_{n\to\infty} P_n^{-1}\phi_n(\theta) = \alpha e^{i\theta} f'(ae^{i\theta}),$$

the limit function obviously being the characteristic function for the distribution whose generating function is $\alpha w f'(aw)$, from which the theorem follows directly.

Now $\mathcal{P}(z)/p$ is the generating function for Z, given $\mathcal{F}$, and the function solves

$$zg(w) - w = 0 \tag{34}$$

for $|z| < \alpha$. We find for the rth moment of Z

$$E(Z^r/\mathcal{F}) = \frac{d^{(r)}}{d(i\theta)^r} \frac{\mathcal{P}(e^{i\theta})}{p}\bigg|_{\theta=0}, \qquad r = 0, 1, \cdots,$$

hence all the moments are finite as soon as $\alpha > 1$. Since by (34)

$$\frac{dw}{dz} = \frac{g(w)}{1 - zg'(w)} = \frac{w}{z(1 - zg'(w))}, \qquad w = \frac{\mathcal{P}(z)}{p},$$

we obtain for the first moment

$$E(Z/\mathcal{F}) = \frac{\mathcal{P}'(1)}{p} = \frac{1}{1 - g'(1)} = \frac{1}{1 - f'(p)}.$$

In a similar way one can express any moment of Z, provided it is finite, in terms of $f'(p)$, $f''(p)$, etc. If $\alpha = 1$ we see from the corollary to theorem 3 that even the first moment of Z is infinite, except for the special case where $\rho = 1$ and $f'(1) < 1$.

The characteristic function of Y_k, given $\mathcal{F}$, is

$$\psi_k(\theta) = 1/p \int_F e^{i\theta Y_k} dP = 1/p \sum_{n=1}^{\infty} \psi_{kn}(\theta),$$

where by (21)

$$\psi_{kn}(\theta) = \int_{\mathfrak{F}_n} e^{i\theta Y_k}\, dP = \sum_{\substack{\Sigma\nu_j = n \\ \Sigma j\nu_j = n-1}} \frac{(n-1)!}{\nu_0\,!\,\nu_1\,!\cdots} e^{i\nu_k\theta} p_0{}^{\nu_0} p_1{}^{\nu_1} \cdots .$$

Thus, if $P_n \neq 0$, $P_n^{-1}\psi_{kn}(\theta)$ is the characteristic function of Y_k, given $\mathfrak{F}_n$. If $p_k = 0$ then $\psi_k(\theta) = 1$. If $p_k \neq 0$ put $p_k = e^{q_k}$ then

$$\frac{\partial^{(r)}}{\partial q_k^r} \mathcal{P}(q) = \sum_{n=1}^{\infty} \psi_{kn}^{(r)}(0) z^n, \tag{35}$$

hence

$$\frac{1}{p} \frac{\partial^{(r)}}{\partial q_k^r} \mathcal{P}(1) = E(Y_k^r/\mathcal{F}),$$

which shows that all moments of Y_k are finite if $\alpha > 1$. Let us put $w = \mathcal{P}(z)$, for short, then, by (18),

$$\frac{\partial w}{\partial q_k} = \frac{z p_k w^k}{1 - z f'(w)} = z^2 p_k w^{k-1} \mathcal{P}'(z), \tag{36}$$

which gives for the first moment of Y_k,

$$E(Y_k/\mathcal{F}) = \frac{p_k p^{k-1}}{1 - f'(p)} = p_k p^{k-1} E(Z/\mathcal{F}),$$

which is to be expected since $p_k p^{k-1}$ plays the same role in $\mathcal{F}$ that p_k plays in $\mathcal{T}$. We may also expect that for $t \in \mathcal{F}_n$, $n^{-1}Y_k$ should be closely related to p_k. This question is settled by the following theorem:

THEOREM 6. *If $a < \rho$ and $q = 1$ then for x real*

$$\lim_{n\to\infty} P(n^{-1}Y_k < x/\mathcal{F}_n) = \begin{cases} 1, & \text{if } x \geq \alpha p_k a^{k-1}; \\ 0, & \text{if } x < \alpha p_k a^{k-1}. \end{cases}$$

PROOF. We intend to estimate the rth moment of $n^{-1}Y_k$ for $t \in \mathcal{F}_n$ and n very large from (35) by means of the contour integral

$$E(n^{-r} Y_k^r/\mathcal{F}_n) = \frac{1}{2\pi i n^r P_n} \int_\Gamma \frac{\partial^{(r)}}{\partial q_k^r} \mathcal{P}(z) \frac{dz}{z^{n+1}}. \tag{37}$$

So let us put

$$w = \mathcal{P}(z), \qquad w_r^{(s)} = \frac{\partial^{(r+s)}}{\partial q_k^r \partial z^s} w, \qquad r, s = 0, 1, \cdots,$$

then by (36) $w_1 = z^2 p_k w^{k-1} w^{(1)}$ and by Leibnitz formula, provided $k \neq 0$,

$$w_r\, z^2 p_k \sum_{\substack{\Sigma\nu_i = r-1 \\ \nu_i \geq 0}} \frac{(r-1)!}{\nu_0!\,\nu_1!\cdots\nu_k!} w_{\nu_1} w_{\nu_2} \cdots w_{\nu_{k-1}} w_{\nu_k}^{(1)}. \tag{38}$$

The principal contribution to the integral in (37) will come from the term of (38) which has the largest size for z near α. If we put $\zeta = (z - \alpha)^{1/2}$ then w is regular at $\zeta = 0$ and so is the constant p_k. Let's assume that w_ν has a pole of order $2\nu - 1$ at $\zeta = 0$ for $\nu = 1, 2, \cdots r - 1$, which is clearly true for $\nu = 1$. Then if s is the number of $\nu_1, \nu_2, \cdots \nu_{k-1}$ which are $= 0$, the order of the pole of the general term of (38) at $\zeta = 0$ is

$$\sum_{i=1}^{k-1} (2\nu_i - 1) + s + 2\nu_k + 1 = 2(r - \nu_0) - (k - s),$$

which has the maximum value $2r - 1$ if and only if $\nu_0 = \nu_1 = \cdots \nu_{k-1} = 0$, $\nu_k = r - 1$. Hence

$$w_r = z^2 p_k w^{k-1} w_{r-1}^{(1)} + \zeta^{2-2r} \mathcal{R}_1(\zeta), \tag{39}$$

where $\mathcal{R}(\zeta)$ is a regular function of ζ at zero. For $k = 0$ the formula (38) is not correct but it is easy to see directly that (39) is correct for $k = 0$. If we derive (39) with respect to z and put $r - 1$ for r we obtain

$$w_{r-1}^{(1)} = z^2 p_k w^{k-1} w_{r-2}^{(2)} + \zeta^{2-2r} \mathcal{R}_2(\zeta),$$

hence

$$w_r = (z^2 p_k w^{k-1})^r w^{(r)} + \zeta^{2-2r} \mathcal{R}_3(\zeta).$$

Substituting in (37) and estimating in a manner similar to that employed previously we obtain

$$\begin{aligned} E(n^{-r} Y_k^r/\mathcal{F}_n) &= \frac{(p_k a^{k-1})^r}{2\pi i n^r p_n} \int_\Gamma \frac{\mathcal{P}^{(r)}(z)}{z^{n-2r+1}}\, dz + \int_\Gamma \frac{(z - \alpha)^{1-r} \mathcal{R}_4((z - \alpha)^{1/2})}{2\pi i n^r p_n z^{n+1}}\, dz \\ &= (p_k a^{k-1})^r \frac{P_{n-r}}{P_n} \frac{(n - r)(n - r - 1) \cdots (n - 2r + 1)}{n^r} + O(n^{-1/2}), \end{aligned}$$

and finally

$$\lim_{n\to\infty} E(n^{-r} Y_k^r/\mathcal{F}_n) = (\alpha p_k a^{k-1})^r.$$

The limit of the rth moment is itself the rth moment of the distribution on the real line which has all its mass at the point $\alpha p_k a^{k-1}$. Since this distribution is uniquely determined by its moments, a well known theorem [7] enables us to conclude that our sequence of distributions has this distribution as limit and this is equivalent to what is claimed by the theorem.

It is important to notice that if we put the mass $\alpha p_k a^{k-1}$ at the point k this determines a distribution on the real line because of (26).

8. The estimation of p. If we wish to estimate p when we know $p \neq 0$, we may obtain an estimate from the knowledge of $f(w)$ in $0 < w < 1$, using the method of iteration. That is we choose a function $G(w)$ such that $G(p) = p$ and $| G(w) - p | < | w - p |$ for $0 < w < 1$. Then if for any w_0 in the open interval we compute successively $w_1, w_2, \cdots$, where $w_{n+1} = G(w_n)$ for $n \geq 0$, we are sure that w_n converges exponentially to p as $n \to \infty$.

Obviously $f(w)$ itself has the properties of $G(w)$ but we achieve faster convergence towards p using Newton's method, that is if we put

$$f_1(w) = f(w) - w,$$

$$G(w) = w - \frac{f_1(w)}{f_1'(w)}. \tag{40}$$

If for some reason we expect p to be close to 1 then it is better to put

$$f_2(w) = \frac{f(w) - w}{w - 1},$$

and use $f_2(w)$ in (40) instead of $f_1(w)$, for then we may choose $w_0 = 1$.

Let us put $f'(1) = 1 + \epsilon$, $\epsilon > 0$ then

$$f_2(1 + h) = \frac{f(1 + h) - 1}{h} - 1 \to \epsilon, \qquad h \to 0;$$

$$f_2'(1 + h) = \lim_{k \to 0} \left(\frac{f(1 + h + k) - 1}{k(h + k)} - \frac{f(1 + h) - 1}{kh} \right),$$

$$f_2'(1) = \lim_{h \to 0} \left(\frac{f(1 + 2h) - 2f(1 + h) + 1}{2h^2} \right) = \frac{f''(1)}{2}.$$

Hence

$$p \approx w_1 = 1 - \frac{2\epsilon}{f''(1)}. \tag{41}$$

This result was previously established by Kolmogoroff [7].

The following two simple examples display the results of the general theory.

EXAMPLE 1. We take $f(w) = p_0 + p_1 w + p_2 w^2$ where $p_0 + p_1 + p_2 = 1$ and $p_0, p_2 > 0$. We have $\rho = \infty$. From the equations (26) and (27),

$$f(a) = p_0 + p_1 a + p_2 a^2 = \frac{a}{\alpha},$$

$$f'(a) = p_1 + 2p_2 a = \frac{1}{\alpha},$$

we obtain easily

$$a = \sqrt{p_0 p_2^{-1}}, \qquad \alpha^{-1} = p_1 + 2\sqrt{p_0 p_2},$$

and it is evident that $a \geq 1$ is equivalent to $p_0 \geq p_2$ is equivalent to $f'(1) = p_1 + 2p_2 \leq 1$. Now

$$\mathcal{G}(z, w) = zp_0 + (zp_1 - 1)w + zp_2 w^2,$$

hence

$$\mathcal{P}(z) = \frac{1 - zp_1 - \sqrt{(1 - zp_1)^2 - 4z^2p_0p_2}}{2zp_2}, \tag{42}$$

the choice of the sign of the radical being determined by letting $z \to 0$.

$$p = \frac{p_0 + p_2 - \sqrt{(p_0 - p_2)^2}}{2p^2} = \begin{cases} 1 , & p_0 \geq p_2 ; \\ p_0 p_2^{-1}, & p_0 < p_2 . \end{cases}$$

In the case $p_1 > 0$ we have $q = 1$ and then by (21)

$$p_n = \sum_{\substack{\nu_0+\nu_1+\nu_2=n \\ \nu_1+2\nu_2=n-1}} \frac{(n-1)!}{\nu_0!\,\nu_1!\,\nu_2!} p_0^{\nu_0} p_1^{\nu_1} p_2^{\nu_2},$$

which can also be obtained by expansion of (42) according to powers of z. From (29) we get

$$p_n \sim \sqrt{\frac{1}{4\pi} (p_1\sqrt{p_0p_2^{-3}} + 2p_0p_2^{-1})}\, (p_1 + 2\sqrt{p_0 p_2})^n\, n^{-3/2}.$$

In the case $p_1 = 0$ we have $q = 2$ and obtain from (42) or from (29)

$$\mathcal{P}(z) = \sum_{\nu=1}^{\infty} (-1)^{\nu+1} \binom{1/2}{\nu} 2^{2\nu-1} p_0^{\nu} p_2^{\nu-1} z^{2\nu-1}$$

$$= \sum_{\nu=1}^{\infty} \frac{(2\nu-2)!}{\nu!\,(\nu-1)!} p_0^{\nu} p_2^{\nu-1} z^{2\nu-1},$$

which shows

$$P_n = \begin{cases} 0 , & n = 2\nu; \\ \dfrac{(2\nu-2)!}{\nu!(\nu-1)!} p_0^{\nu} p_2^{\nu-1}, & n = 2\nu - 1. \end{cases}$$

By direct use of Stirling's formula or from (29) we get

$$P_{2\nu-1} \sim \frac{1}{p_2} \sqrt{\frac{2}{\pi}}\, 2^{2\nu-1} (p_0p_2)^{\nu} (2\nu-1)^{3/2}.$$

EXAMPLE 2. We take $f(w) = e^{\lambda(w-1)}$, $\lambda > 0$, so that W has a Poisson distribution. Then $\rho = \infty$, $q = 1$, and we get from (26) and (27)

$$f(a) = e^{\lambda(a-1)} = a/\alpha,$$

$$f'(a) = \lambda e^{\lambda(a-1)} = 1/\alpha,$$

$$a = 1/\lambda, \qquad \alpha = e^{\lambda-1}/\lambda.$$

Clearly we have $a \geq 1$ if and only if $\lambda \leq 1$ and in this case 1 is evidently the only solution for w of $e^{\lambda(w-1)} = w$, hence $p = 1$. On the other hand if $\lambda < 1$

then (41) gives $p = 1 - 2(\lambda - 1)\lambda^{-2}$. By (21) we get

$$P_n = \frac{(n\lambda)^{n-1}}{n!} e^{-n\lambda},$$

and by direct use of Stirling's formula or from (29) we get

$$P_n \sim \sqrt{\frac{1}{2\pi}}\, e^{n(1-\lambda)} \lambda^{n-1} n^{-3/2}.$$

REFERENCES

[1] R. A. FISHER, *The Genetical Theory of Natural Selection*, Oxford, The Clarendon Press 1930.

[2] A. J. LOTKA, *Theorie Analytique des Associations Biologiques* 2, Hermann and Co., Paris, 1939.

[3] D. HAWKINS AND S. ULAM, *Theory of Multiplicative Processes* I, MDDC-287, 1944.

[4] T. E. HARRIS, "Some theorems on the Bernoullian multiplicative process," *thesis, doctor of philosophy*, Princeton University, 1947.

[5] A. M. YAGLOM, "Certain limit theorems of the theory of branching random processes," *Doklady Akad. Nauk. SSSR(N. S.)* Vol. 56 (1947), 795–798.

[6] D. G. KENDALL, "On the generalized "birth-and-death" process," *Annals of Math. Stat.*, Vol. 19 (1948).

[7] A. KOLMOGOROFF, "Zur Lösung einer biologischen Aufgabe," *Mitt. Forsch.- Inst. Math. u. Mech. Univ. Tomsk*, Vol. 2 (1938), pp. 1–6.

[8] L. PONTRJAGIN, *Topological Groups*, Princeton Univ. Press, 1946.

[9] J. VON NEUMANN, *Functional Operators*, mimeographed notes, Institute for Advanced Study, 1933–35.

[10] A. KOLMOGOROFF, *Grundbergriffe der Wahrscheinlichkeitsrechnung*, Chelsea Publishing Co., New York, 1946.

[11] A. HURWITZ AND R. COURANT, *Funktionentheorie*, Springer, Berlin, 1929.

[12] M. G. KENDALL, *The Advanced Theory of Statistics*, Vol. I, Griffin Co., London, 1943.