THE GENERAL THEORY OF STOCHASTIC POPULATION PROCESSES

BY

J. E. MOYAL

Australian National University, Canberra, Australia (¹)

1. Introduction

The purpose of the present paper is to lay down the foundations of a general theory of stochastic population processes (see Bartlett [2] for references to previous work on this subject). By *population* we mean here a collection of individuals, each of which may be found in any one state x of a fixed set of states X. The simplest type of population process is one where there is only one kind of individual and where the total size of the population is always finite with probability unity (finite univariate population process). The state of the whole population is characterized by the states, say $x_1, \dots, x_n$, of its members, where each of the x_i ranges over X and $n = 0, 1, 2, \dots$; thus we may have for example a biological population whose individuals are characterized by their age, weight, location, etc. or a population of stars characterized by their brightness, mass, position, velocity and so on. Such a population is *stochastic* in the sense that there is defined a probability distribution P on some σ-field $\mathbf{B}$ of subsets of the space $\mathfrak{X}$ of all population states; in § 2 we develop the theory of such *population probability spaces*; the approach is similar to that of Bhabha [4], who, however, restricts himself to the case where X is the real line and P is absolutely continuous. By taking the individual state space X to be arbitrary (i.e., an abstract space), we are able to make the theory completely general, including for example the case of cluster processes, where the members of a given population are clusters, i.e., are themselves populations, as in Neyman's theory of populations of galaxy clusters (cf. Neyman

(¹) This work was supported in part by Office of Naval Research Contract Nonr-225(21) at Stanford University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

and Scott [13]). In § 3 we develop an alternative and largely equivalent approach where the state of the population is characterized by an integral-valued *function* N on a class of subsets of $X, N(A)$ representing the number of individuals in the population with states in the subset A of X. § 4 is devoted to the study of *generating functionals,* which play a role analogous to that of probability and moment generating *functions* in standard probability theory; generating functionals were introduced in this connection in Kendall [9] and Bartlett and Kendall [3] (see also Bartlett [1]). In § 5 we extend the theory to *multivariate populations,* where there is more than one kind of individual (e.g., biological populations comprising several species) and *population processes,* where the population state is a function of some independent variable such as time or space coordinates: an example is that of cosmic ray cascades, which are populations of several kinds of elementary particles (electrons, photons, nucleons, mesons, etc.), characterized by their energy, momentum, position, etc., developing in time through the atmosphere. In § 6 we extend the theory still further to the case of populations whose total size can be infinite with positive probability. Finally, in § 7 we consider as examples special types of population processes: cluster processes; processes with independent elements; Markov population processes, where we treat the problem of obtaining the probability distribution of the process from given reproduction and mortality rates; lastly, the important case of multiplicative population processes, where each "ancestor" generates a population independent of the "descendents" of other "ancestors".

2. Point Processes

We start with the space X of all individual states x. If the members of the population can be distinguished from each other, then a *state of the population* is defined as an ordered set $x^n = (x_1, \ldots, x_n)$ of individual states: i.e., it is the state where the population has n members, the ith being in the individual state x_i, $i = 1, \ldots, n$. The *population state space* $\mathfrak{X}$ is the class of all such states x^n with $n = 0, 1, 2, \ldots, x^0$ denoting conventionally the state where the population is empty.

If the members of the population are indistinguishable from each other, then a state of the population is defined as an *unordered set* $x^n = \{x_1 \ldots, x_n\}$ of individual states: i.e., a state where the population has n members with one each in the states $x_1, \ldots, x_n$. The population state space $\mathfrak{X}$ is now the set of all such x^n with $n = 0, 1, 2, \ldots, x^0$ denoting again an empty population.

A triplet $(\mathfrak{X}, \boldsymbol{B}, P)$, where $\mathfrak{X}$ is a population state space, $\boldsymbol{B}$ a σ-field of sets in $\mathfrak{X}$ and P a probability distribution on $\boldsymbol{B}$ constitutes a model of a stochastic population and will be called a *point process*, (1) or more precisely a single-variate point process. We shall often use the single letter x to symbolize a point process.

Let X^n be the set of all states x^n with n fixed; X^n is a subset of $\mathfrak{X}$. If the individuals are distinguishable, then X^n is the Cartesian product $X \times X \times \ldots \times X$, n times. If $n \neq k$, then X^n and X^k are disjoint, and all subsets $A^{(n)}$, $A^{(k)}$ of respectively X^n and X^k are disjoint. It follows that every set A in $\mathfrak{X}$ can be expressed uniquely as a sum of disjoint sets $A = \sum_{n=0}^{\infty} A^{(n)}$, where $A^{(n)} = A \cap X^n$ (throughout this paper we shall use $\cap$, $\cup$ to denote respectively set intersections and unions, and $+$, Σ to denote unions of mutually disjoint sets). The following results are obvious: let $A, A_1, A_2, A_t (t \in T)$ be sets in $\mathfrak{X}$, $-A$ the complement of A, $\varnothing$ the empty set and $A^{(n)} = A \cap X^n$; then

$$-A = \sum_{n=0}^{\infty} (X^n - A^{(n)}), \tag{2.1}$$

$$A_1 \cap A_2 = \varnothing \text{ implies that } A_1^{(n)} \cap A_2^{(n)} = \varnothing \quad (n = 0, 1, 2, \ldots), \tag{2.2}$$

$$\bigcup_{t \in T} A_t = \sum_{n=0}^{\infty} \bigcup_{t \in T} A_t^{(n)} \quad \text{and} \quad \bigcap_{t \in T} A_t = \sum_{n=0}^{\infty} \bigcap_{t \in T} A_t^{(n)}. \tag{2.3}$$

If we wish to differentiate between the two cases of distinguishable and indistinguishable individuals, we shall denote the population state space in the latter case by $\mathfrak{X}_S$, points in $\mathfrak{X}_S$ by $x^{\{n\}}$, the set of all points $x^{\{n\}}$ with n fixed by $X^{\{n\}}$, subsets of $X^{\{n\}}$ by $A^{\{n\}}$, and so on. Let π be the permutation $i_1, \ldots, i_n$ of $1, \ldots, n$, let $A_\pi^{(n)}$ be the subset of X^n obtained from $A^{(n)}$ by the permutation $(x_{i_1}, \ldots, x_{i_n})$ of the *coordinates* $x_1, \ldots, x_n$ of each x^n in $A^{(n)}$. A in $\mathfrak{X}$ is *symmetric* if $A^{(n)} = A \cap X^n = A_\pi^{(n)}$ for all permutations π. The *symmetrization* of A is the symmetric set $\bigcup_\pi A_\pi$. The relation between the two cases of distinguishable and indistinguishable individuals is that given X, we have the transformation T from $\mathfrak{X}$ onto $\mathfrak{X}_S$ which maps every $x^n = (x_1, \ldots, x_n)$ into $x^{\{n\}} = \{x_1, \ldots, x_n\}$; clearly the inverse image $T^{-1}\{x^{\{n\}}\} = \sum_\pi \{x^n\}_\pi$, and hence there is a one-to-one correspondence between subsets of $\mathfrak{X}_S$ and *symmetric* subsets of $\mathfrak{X}$: thus we can *identify* sets of states in a population with indistinguishable individuals with symmetric subsets of the state space $\mathfrak{X}$ of a population with distinguishable individuals.

Let $\mathbf{B}$ be a σ-field of sets in X; we call the pair $(X, \mathbf{B})$ the *individual measure space*. Let $\mathbf{B}^n$ be the minimal σ-field of sets in X^n cotaining all product sets $A_1 \times \ldots \times A_n$ such that $A_i \in \mathbf{B}$, $i = 1, \ldots, n$, and let $\boldsymbol{\mathcal{B}}$ be the class of all sets $A = \sum_{n=0}^{\infty} A^{(n)}$ in $\mathfrak{X}$ such that $A^{(n)} \in \mathbf{B}^n$, $n = 0, 1, 2, \ldots$; by an abuse of standard notation we write $\boldsymbol{\mathcal{B}} = \sum_{n=0}^{\infty} \mathbf{B}^n$; then

(1) The term "point process" is due to Wold [17]: see also Bartlett [1] and [2].

LEMMA 2.1. *$\mathcal{B}$ is the minimal σ-field of sets in $\mathfrak{X}$ containing all sets $A^{(n)} \in \mathbf{B}^n$, $n = 0, 1, 2, \ldots$*

Proof. That $\mathcal{B}$ is a σ-field follows immediately from (2.1) and (2.3); if $\mathcal{B}'$ is a σ-field containing all $A^{(n)} \in \mathbf{B}^n$, it will contain all unions $\sum_{n=0}^{\infty} A^{(n)}$ of such $A^{(n)}$; hence $\mathcal{B}' \supset \mathcal{B}$.

We call the pair $(\mathfrak{X}, \mathcal{B})$ the *population measure space*. In the case of a population with indistinguishable individuals, starting from $(X, \mathbf{B})$, we define $(\mathfrak{X}, \mathcal{B})$ as above, and then define the σ-field $\mathcal{B}_S$ to be the class of all sets A in $\mathfrak{X}_S$ such that $T^{-1}A \in \mathcal{B}$; thus $\mathcal{B}_S$ may be identified with the σ-subfield of all *symmetric* sets of $\mathcal{B}$. Sets of $\mathcal{B}$ or $\mathcal{B}_S$ will be called *measurable*. In order to avoid trivial difficulties, we shall assume as a rule that $\mathbf{B}$ contains all singletons $\{x\}$; it then follows that $\mathcal{B}$ contains all singletons $\{x^n\}$ and $\mathcal{B}_S$ contains all singletons $\{x^{\{n\}}\}$.

We now turn to the definition of a probability distribution P on $\mathcal{B}$. The restriction $P^{(n)}$ of P to sets in X^n is a measure on $\mathbf{B}^n$, $p_n = P^{(n)}(X^n)$ is the probability that the size of the population is n, and $\sum_{n=0}^{\infty} p_n = 1$. Conversely, we have:

LEMMA 2.2. *If $P^{(n)}$ is a measure on $\mathbf{B}^n$, $n = 0, 1, \ldots$ such that $\sum_{n=0}^{\infty} P^{(n)}(X^n) = 1$, then the function P on $\mathcal{B}$ whose value at $A = \sum_{n=0}^{\infty} A^{(n)}$ is*

$$P(A) = \sum_{n=0}^{\infty} P^{(n)}(A^{(n)}) \tag{2.4}$$

is the unique probability distribution on $\mathcal{B}$ whose restriction to $\mathbf{B}^n$ agrees with $P^{(n)}$ for all n.

Proof. Clearly $P \geqslant 0$, $P(\mathfrak{X}) = \sum_{n=0}^{\infty} P^{(n)}(X^n) = 1$; if $\{A_k\}$ is a disjoint sequence of measurable sets, then by (2.3) and the definition of P

$$P\left(\sum_{k=1}^{\infty} A_k\right) = P\left(\sum_{n=0}^{\infty} \sum_{k=1}^{\infty} A_k^{(n)}\right) = \sum_{n=0}^{\infty} \sum_{k=1}^{\infty} P^{(n)}(A_k^{(n)}) = \sum_{k=1}^{\infty} P(A_k);$$

hence P is σ-additive and therefore a probability distribution on $\mathcal{B}$. Finally, if P' is a probability distribution on $\mathcal{B}$ which agrees with $P^{(n)}$ on $\mathbf{B}^n$ for all n, then clearly $P'(A) = P(A)$ for every $A \in \mathcal{B}$.

Exactly the same considerations apply to a probability distribution P_S on the σ-field $\mathcal{B}_S$ in the case of indistinguishable individuals. A probability distribution P on $\mathcal{B}$ determines a unique probability distribution P_S on $\mathcal{B}_S$ such that for every $A \in \mathcal{B}_S$

$$P_S(A) = P(T^{-1}A). \tag{2.5}$$

Conversely, it can be shown that a probability distribution P_S on $\mathbf{B}_S$ determines a unique *symmetric* probability distribution P on $\mathbf{B}$ (by this we mean a distribution on $\mathbf{B}$ which is invariant under coordinate permutations) to which it is related by (2.5). Thus we can identify a distribution on $\mathbf{B}_S$ with the corresponding symmetric distribution on $\mathbf{B}$, and we shall use the same symbol P_S for both. It follows from the foregoing discussion that a point process $(\mathfrak{X}, \mathbf{B}, P_S)$ with distinguishable individuals and a symmetric distribution and the corresponding process $(\mathfrak{X}_S, \mathbf{B}_S, P_S)$ with indistinguishable individuals are to all intents and purposes the same thing; we shall henceforth identify the two under the name of *symmetric point process* (symbolized by the single letter $\mathbf{x}_S$). With each point process $(\mathfrak{X}, \mathbf{B}, P)$ is associated a unique symmetric point process $(\mathfrak{X}, \mathbf{B}, P_S)$ obtained by *symmetrization* of P: that is for each $A \in \mathbf{B}$,

$$P_S(A) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{\pi} P^{(n)}(A_\pi^{(n)}). \tag{2.6}$$

Point processes and their distributions which are associated with the same symmetric point process will be called *countably equivalent.*

A particularly simple class of point processes is the class of *compound processes* (cf. Feller [6], p. 268) where given that the population is of size n, the states of the n individuals are independently distributed with the same distribution Q on $\mathbf{B}$; hence

$$P^{(n)} = p_n Q^{*n} \quad (n = 0, 1, 2, \ldots), \tag{2.7}$$

where p_n is the probability that the population is of size n, Q^{*n} is the nth product measure on $\mathbf{B}^n$ generated by Q; clearly $P^{(n)}$ is symmetric. Two examples are the *Poisson process,* where $p_n = m^n e^{-m}/n!\ (m > 0)$ and the *geometric process,* where $p_n = (1-q)q^n$ $(0 < q < 1)$.

A real-valued function y on $\mathfrak{X}$ is a mapping of $\mathfrak{X}$ into the real line; y may be identified with the sequence $\{y_n\}$ where y_n is the restriction of y to X^n, and is therefore a real-valued function on X^n. We say that y is *measurable* if $y^{-1}(S) \in \mathbf{B}$ for every Borel subset S of the real line; similarly y_n is measurable if $y_n^{-1}(S) \in \mathbf{B}^n$ for every Borel set S. The following lemma is evident:

LEMMA 2.3. *A function y on $\mathfrak{X}$ is measurable if and only if each of its restrictions y_n is a measurable function on X^n, $n = 0, 1, 2, \ldots$.*

It follows that a sequence of measurable functions y_n on X^n, $n = 0, 1, 2, \ldots$, determines a unique measurable function y on $\mathfrak{X}$ whose restrictions to X^n is y_n. Following the usual terminology, we say that a measurable function y on $\mathfrak{X}$ determines a *random* *variable* on the point process $(\mathfrak{X}, \mathbf{B}, P)$. The *expectation value* of y with respect to P, defined in the usual way,

$$Ey = \int_{\mathfrak{X}} y(x^n)\, P(dx^n) = \sum_{n=0}^{\infty} \int_{X^n} y_n(x^n)\, P^{(n)}(dx^n), \tag{2.8}$$

clearly exists if and only if each integral $\int_{X^n} y_n P^{(n)}(dx^n)$ in the series above exists and their sum is absolutely convergent. If y is *symmetric* (i.e., invariant under coordinate permutations) then $\int_{\mathfrak{X}} y\, P(dx^n) = \int_{\mathfrak{X}} y\, P_S(dx^n)$, where P_S is the symmetrization of P; conversely, if P is symmetric, then $\int_{\mathfrak{X}} y\, P(dx^n) = \int_{\mathfrak{X}} y^{(S)} P(dx^n)$, where $y^{(S)}$ is the *symmetrization* of y:

$$y_n^{(S)}(x^n) = \frac{1}{n!} \sum_{\pi} y_n(x_\pi^n), \quad (n = 0, 1, 2, \ldots). \tag{2.9}$$

It will be convenient to admit infinite expectations: i.e., $Ey = \pm\infty$ if the right-hand side of (2.8) diverges definitely to $\pm\infty$.

3. Counting Processes

We shall now turn our attention to an alternative method of characterizing stochastic populations, namely, the method which assigns to sets A in the individual state space X the number of individuals $N(A)$ which are in states $x \in A$; *we restrict ourselves in the present section to populations whose total size $N(X)$ is finite with probability unity.* It is convenient to treat this approach by relating it to the previous one. It is intuitively obvious that if in a population of total size n there are one individual each in the states $x_1, \ldots, x_n$, then the number of individuals with states in a given arbitrary subset A of X is given by the expression

$$N(A \mid x^{\{n\}}) = \sum_{i=1}^{n} \delta(A \mid x_i): \tag{3.1}$$

where $\delta(A \mid x)$ is the characteristic function of the set A:

$$\delta(A \mid x) = \begin{cases} 1 & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A. \end{cases} \tag{3.2}$$

For each fixed A in X, $N(A \mid \cdot)$ is a function on the symmetric population state space $\mathfrak{X}_S$, while for each fixed $x^{\{n\}} \in \mathfrak{X}_S$, $N(\cdot \mid x^{\{n\}})$ is a function on the class $\mathbf{U}$ of all subsets of X. Clearly each such function on $\mathbf{U}$ will have the following properties: it is (1) non-negative, (2) finite, (3) integral-valued and (4) *completely additive,* in the sense that if

$\{A_t, t \in T\}$ is an arbitrary indexed collection of mutually disjoint sets in X, then at most a finite number of sets, say $A_{t_1}, \ldots, A_{t_n}$, can be such that $N(A_t) \geqslant 1$ and $N(\sum_{t \in T} A_t) = \sum_{t \in T} (N|At) = \sum_{i=1}^n N(A_{t_i})$. The converse is also true; call *counting measure* a function on $\mathbf{U}$ which has the properties (1) to (4) above, and let $\mathcal{N}$ be the *class of all counting measures on* $\mathbf{U}$; then

THEOREM 3.1. *Relation* (3.1) *defines a one-to-one correspondence between* $\mathfrak{X}_S$ *and* $\mathcal{N}$.

Proof. It follows immediately from the additivity of N that $N(\emptyset) = 0$. Let $N(X) = n$, and let k_1 be the smallest integer such that there exists a set A_1 such that $N(A_1) = k_1$: then $k_1 \leqslant n$ and $N(X - A_1) = n - k_1$. If $k_1 < n$, let k_2 be the smallest integer such that there exists a subset A_2 of $X - A_1$ such that $N(A_2) = k_2$: then $k_1 + k_2 \leqslant n$ and $N(X - (A_1 + A_2)) = n - k_1 - k_2$. Repetition of this procedure must obviously terminate after $r \leqslant n$ step, yielding r pairs (k_i, A_i) such that $\sum_{i=1}^r k_i = n, A_i \cap A_j = \emptyset$ if $i \neq j$ and $\sum_{i=1}^r A_i = X$. It follows that if $B \subset A_j$, then either $N(B) = k_j$ and $N(A_j - B) = 0$ or $N(B) = 0$ and $N(A_j - B) = k_j$. Let $\mathbf{G}$ be the class of all subsets B of A_j such that $N(B) = 0$. Let us index the elements of $\mathbf{G}$ with the indices $t \in T$; we may assume that T is the class of all ordinals $t < \alpha$, where α is the least ordinal with the same power as T: thus $\mathbf{G} = \{B_t, t < \alpha\}$. We now show that there exists a disjoint indexed class $\{C_t, t < \alpha\}$ of subsets of A_j such that $B_\alpha = \bigcup_{t<\alpha} B_t = \sum_{t<\alpha} C_t$. Let $C_0 = B_0$ and let $C_t = B_t - \bigcup_{s<t} B_s$; then clearly $C_t \cap C_{t'} = \emptyset$ if $t \neq t'$. Suppose that $\bigcup_{s<t} B_s = \sum_{s<t} C_s$ for some $t > 0$ and $\leqslant \alpha$. Then clearly this equality holds for $t+1$. Furthermore if t is a limit ordinal and the equality holds for all $s < t$, then $B_s \subset \sum_{r \leqslant s} C_r$ and hence $\bigcup_{s<t} B_s \subset \sum_{s<t} C_s$. On the other hand from the definition $C_t \subset B_t$ for every $t < \alpha$ and hence $\sum_{s<t} C_s \subset \bigcup_{s<t} B_s$; hence $\bigcup_{s<t} B_s = \sum_{s<t} C_s$. It follows by transfinite induction that $\sum_{t<\alpha} C_t = \bigcup_{t<\alpha} B_t$. Therefore by the assumed complete additivity of N,

$$N(B_\alpha) = N(\bigcup_{t<\alpha} B_t) = N(\sum_{t<\alpha} C_t) = \sum_{t<\alpha} N(C_t) = 0,$$

since $C_t \subset B_t$ and $N(B_t) = 0$ for all $t < \alpha$. Hence if we let $D = A_j - B_\alpha$, then $N(D) = k_j$; therefore $D \neq \emptyset$. We assert that D is a singleton, say $\{y_j\}$; for suppose not; then there exists a non-empty proper subset E of D such that $N(E) = 0$; hence $E \in \mathbf{G}$ and therefore $D - E = (A_j - B_\alpha) - E = A_j - B_\alpha = D$; hence $E = \emptyset$, which is a contradiction; hence $D = \{y_j\}$, $N\{y_j\} = k_j$, and for every $B \subset A_j$, $N(B) = k_j$ if $y_j \in B$ and $N(B) = 0$ if $y_j \notin B$. This result is true for $j = 1, 2, \ldots, r$. Hence if we set $x^{(n)} = \{x_1, \ldots, x_n\}$, where $x_1 = x_2 = \ldots = x_{k_1} = y_1$, $x_{k_1+1} = \ldots = x_{k_1+k_2} = y_2$ and so on, then for every $A \subset X$, $N(A) = \sum_{i=1}^n \delta(A|x_i)$. This completes the proof that relation (3.1) defines a one-to-one correspondence between $\mathfrak{X}_S$ and $\mathcal{N}$; we shall denote this mapping by the letter T_{SN}.

If now we are given the symmetric point process $(\mathfrak{X}_S, \mathbf{B}_S, P_S)$, let $\mathbf{B}_N = T_{SN}^{-1}\mathbf{B}_S$, i.e., $\mathbf{B}_N$ is the σ-field of all sets of functions $N \in \mathcal{N}$ whose inverse image belongs to $\mathbf{B}_S$. Let P_N be the probability distribution on $\mathbf{B}_N$ such that $P_N(A) = P_S(T_{SN}^{-1}A)$ for every $A \in \mathbf{B}_N$; we write $P_N = T_{SN}^* P_S$. Then *the triplet* $(\mathcal{N}, \mathbf{B}_N, P_N)$ *constitutes a probability space, which we shall call a counting process* and denote briefly by $\mathbf{N}$. In other words, T_{SN} defines a one-to-one measure-preserving transformation from $\mathbf{x}_S$ onto $\mathbf{N}$.

We now turn to a closer study of the σ-field $\mathbf{B}_N$. We note that for fixed A, the function $N(A \mid \cdot)$ on $\mathfrak{X}_S$ defined by (3.1) is measurable if and only if A is measurable. This leads to the following theorem:

THEOREM 3.2. *The σ-field $\mathbf{B}_N$ is the smallest σ-field containing all sets $\{N \mid N(A_i) = k_i;\ i = 1, 2, \ldots, n\}$ in $\mathcal{N}$, where A_i is measurable and k_i is a non-negative integer, $i = 1, \ldots, n$.*

Proof. Let $\{A_1, \ldots, A_r\}$ be a measurable finite partition of X: i.e., a disjoint finite collection of sets in X such that $\sum_{i=1}^r A_i = X$. Let $\mathbf{C}_S$ denote the class of all symmetrized product sets

$$(A_1^{k_1} \times \ldots \times A_r^{k_r})_S = \sum_{\pi} (A_1^{k_1} \times \ldots \times A_r^{k_r})_\pi \tag{3.3}$$

formed from such partitions for all finite sets of non-negative integers $\{k_1, \ldots, k_r\}$ with $r = 1, 2, \ldots,$ and let $\mathbf{C}_N$ be the class of all sets $\{N \mid N(A_i) = k_i;\ i = 1, \ldots, r\}$. The transformation T_{SN} establishes a one-to-one correspondence between elements of $\mathbf{C}_S$ and $\mathbf{C}_N$:

$$T_{SN}(A_1^{k_1} \times \ldots \times A_r^{k_r})_S = \{N \mid N(A_i) = k_i;\ i = 1, \ldots, r\}, \tag{3.4}$$

since by Theorem 3.1 if $n = k_1 + \ldots + k_r$, then each function N satisfying the condition in the right-hand side of (3.4) determines a unique $x^{\{n\}} = \{x_1, \ldots, x_n\}$ such that k_i of the coordinates belong to A_i, $i = 1, \ldots, r$, and conversely. We now remark that if $\{B_1, \ldots, B_n\}$ is an arbitrary finite collection of non-empty measurable sets in X, then there exists a finite measurable partition $\{A_1, \ldots, A_r\}$ of X for some $r \geqslant n$ such that each B_j, $j = 1, \ldots, n$ is the union of one or more of the A_i. It follows that the symmetrized product set $(B_1 \times \ldots \times B_n)_S$ is the union of a disjoint collection of sets of the form $(A_{i_1}^{k_1} \times \ldots \times A_{i_h}^{k_h})$, where $h \geqslant 1$ and $k_1 + \ldots + k_h = n$, i.e., of sets belonging to $\mathbf{C}_S$. Now by Lemma 2.1 $\mathbf{B}_S$ is the minimal σ-field containing all symmetrized product sets $(B_1 \times \ldots \times B_n)_S$, and hence is the σ-field generated by $\mathbf{C}_S$. It then follows from (3.4) and Theorem 3.1 that $\mathbf{B}_N$ is the σ-field generated by $\mathbf{C}_N$. Because of the additivity of N, the random variables $N(B_j)$, $j = 1, \ldots, n$, can be expressed as sums of one or more of the random variables $N(A_i)$, and hence each set of the form $\{N \mid N(B_j) = k_j;\ j = 1, \ldots, n\}$ can be expressed as a union of sets belonging to $\mathbf{C}_N$. Hence the smallest

σ-field containing all such sets coincides with the σ-field generated by $\mathbf{C}_N$, i.e., with $\boldsymbol{B}_N$. This completes the proof of Theorem 3.2.

It follows from (3.4) that if $\{A_1, \ldots, A_r\}$ is a finite measurable partition of X and $k_1 + \ldots + k_r = n$ then

$$P_N\{N(A_i) = k_i;\ i = 1, \ldots, r\} = P_S^{(n)}(A_1^{k_1} \times \ldots \times A_r^{k_r})_S$$
$$= \frac{n!}{k_1! \ldots k_r!} P_S^{(n)}(A_1^{k_1} \times \ldots \times A_r^{k_r}), \tag{3.5}$$

where P_S is a symmetric distribution on $\boldsymbol{B}$. It will be seen from the proof of Theorem 3.2 that we can use (3.5) in order to express the joint distributions of $N(A)$ for any finite collection of measurable sets in X in terms of the distribution P_S on $\boldsymbol{B}$. Thus for example the distribution of $N(A)$ is

$$P_N(N(A) = n) = \sum_{k=0}^{\infty} P_N(N(A) = n,\ N(X - A) = k)$$
$$= \sum_{k=0}^{\infty} \binom{n+k}{n} P_S^{(n+k)}(A^n \times (X - A)^k). \tag{3.6}$$

In the case of a Poisson process, where $P_S^{(n)} = m^n e^{-m} Q^{*n}/n!$,

$$P_N\{N(A_i) = k_i;\ i = 1, \ldots, r\} = \frac{n!}{k_1! \ldots k_r!} \frac{m^n}{n!} e^{-m} Q^{k_1}(A_1) \ldots Q^{k_r}(A_r)$$
$$= \prod_{i=1}^{r} \frac{m^{k_i} Q^{k_i}(A_i)}{k_i!} e^{-mQ(A_i)}; \tag{3.7}$$

hence the $N(A_i)$ are mutually independent Poisson variates with means $mQ(A_i)$, $i = 1, \ldots, r$.

Theorem 3.2 leads to the conjecture that in the definition of a counting process we could restrict the domain of the functions N to the σ-field $\mathbf{B}$. This is in fact true if $\mathbf{B}$ includes all singletons, in the sense that this last condition together with conditions (1) to (4) in the definition of $\mathcal{N}$ ensure the one-to-one correspondence between $\mathfrak{X}_S$ and $\mathcal{N}$; the proof of Lemma 3.1 can be modified to show this. We take in the proof all the sets A_j, $j = 1, \ldots, r$ and B_t, $t \in T$ to be measurable; then $B_\alpha = \bigcup_{t<\alpha} B_t$ is measurable. For if it is not then $B_\alpha \neq A_j$; hence $A_j - B_\alpha \neq \emptyset$ and therefore contains at least one point y; but $\{y\} \in \mathbf{B}$ by hypothesis; therefore the smallest measurable set B_0 which contains B_α is a proper subset of A_j. Suppose $N(B_0) \neq 0$; then $N(A_j - B_0) = 0$; hence $A_j - B_0 \in \boldsymbol{G}$, which is a contradiction; hence $N(B_0) = 0$ and therefore $B_0 = B_\alpha$;

hence B_α is measurable. The rest of the proof follows as before. We see thus that the restriction is in fact trivial, since under these conditions each N defined on $\mathbf{B}$ extends uniquely to the σ-field $\mathbf{U}$ of all subsets of X. It will be convenient nevertheless in what follows to make this restriction, since we shall always be dealing with measurable sets only. An open question is whether σ-additivity alone is sufficient to ensure the truth of Theorem 3.1: i.e., does σ-additivity imply complete additivity for finite integral-valued measures on $\mathbf{U}$ or $\mathbf{B}$? Ulam (1) [16] has shown that a sufficient additional condition for such measures on $\mathbf{U}$ is that *the cardinal of X be accessible*, and one may conjecture that the same is true for such measures on a σ-field $\mathbf{B}$ containing all singletons.

We shall now consider the *measures* generated by the moments of a counting process $\mathbf{N}$. Writing as before $p_n = P^{(n)}(X^n)$, let $m_k = \sum_1^\infty n^k p_n$ and $m_{(k)} = \sum_k^\infty n!\,p_n/k!$ be respectively the kth moment and factorial moment of the distribution $\{p_n\}$, and write $m_1 = m_{(1)} = m$.

LEMMA 3.3. *The mean*

$$M(A) = EN(A) = \sum_{n=1}^{\infty} \int_{X^n} \sum_{i=1}^{n} \delta(A\,|\,x_i)\, P_S^{(n)}(dx^n) = \sum_{n=1}^{\infty} n\, P_S^{(n)}(A \times X^{n-1}) \tag{3.8}$$

defines the value at A of a measure M on $\mathbf{B}$ which is finite if and only if $m < \infty$, and σ-finite if and only if $M(A_i) < \infty$ for each A_i of some measurable countable partition $\{A_i\}$ of X.

Proof. The lemma follows at once by known results from the fact that M is seen in the right-hand side of (3.8) to be the sum of a series of measures on $\mathbf{B}$. We shall say in what follows that M *exists* only if it exists as a finite or σ-finite measure.

We shall call M the *mean distribution* generated by N. It is instructive to rederive (3.8) using expression (3.6) for $P_N(N(A) = n)$ and the relation

$$\begin{aligned} P_S^{(k)}(A \times X^{k-1}) &= \int_{X^k} \delta(A\,|\,x_i) \prod_{i=2}^{k} [\delta(A\,|\,x_i) + \delta(X - A\,|\,x_i)]\, P_S^{(k)}(dx^n) \\ &= \sum_{j=0}^{k-1} \binom{k-1}{j} P_S^{(k)}(A^{j+1} \times (X-A)^{k-j-1}). \end{aligned} \tag{3.9}$$

From (3.6) and (3.9) we obtain

(1) I am indebted to Drs. Erdös and Le Cam for this reference.

$$M(A)=\sum_{n=0}^{\infty} n P_N(N(A)=n)=\sum_{n=0}^{\infty} n \sum_{k=n}^{\infty}\binom{k}{n} P_S^{(k)}(A^n\times(X-A)^{k-n})$$

$$=\sum_{k=1}^{\infty} k \sum_{j=0}^{k-1}\binom{k-1}{j} P_S^{(k)}(A^{j+1}\times(X-A)^{k-j-1})=\sum_{k=1}^{\infty} k P_S^{(k)}(A\times X^{k-1}).$$

Consider now the kth product measure N_k generated by N on $\mathbf{B}^k$. It follows from (3.1) that for each $A^{(k)}\in\mathbf{B}^k$

$$N_k(A^{(k)}\,|\,x^{\{n\}})=\sum_{i_1=1}^{n}\dots\sum_{i_k=1}^{n}\delta(A^{(k)}\,|\,x_{i_1},\dots,x_{i_k}) \tag{3.10}$$

where $\delta(\cdot\,|\,x_1,\dots,x_k)$ is the product measure $\delta(\cdot\,|\,x_1)\times\dots\times\delta(\cdot\,|\,x_k)$; in other words, the function $N_k(\cdot\,|\,x^{\{n\}})$ assigns measure 1 to every singleton $\{(x_{i_1},\dots,x_{i_k})\}$ in X^k such that $x_{i_j}\in\{x_1,\dots,x_n\}$, $j=1,\dots,k$ and measure 0 to the rest of X^k. Hence N_k *is itself a counting measure on* $\mathbf{B}^k$ *and the expectation* $M_k=EN_k$ *is the mean distribution of* N_k; it therefore follows from Lemma 3.3 that M_k *is a measure on* $\mathbf{B}^k$, finite if and only if $m_k<\infty$, σ-finite if and only if $M_k(A_i^{(k)})<\infty$ for each set $A_i^{(k)}$ of some measurable countable partition of X^k. We call M_k the *k-th moment distribution* generated by N.

The expression of M_k in terms of P_S is rather complicated in general. A much simpler expression obtains for the *k-th factorial moment distribution* $M_{(k)}=EN_{(k)}$, where $N_{(k)}$ is the counting measure on $\mathbf{B}^k$ defined by the relation

$$N_{(k)}(A^{(k)}\,|\,x^{\{n\}})=\sum_{i_1\neq i_2\neq\dots\neq i_k}\delta(A^{(k)}\,|\,x_{i_1},\dots,x_{i_k})\quad(n\geqslant k); \tag{3.11}$$

in other words $N_{(k)}(\cdot\,|\,x^{\{n\}})$ assigns measure one to every singleton $\{(x_{i_1},\dots,x_{i_k})\}$ such that $x_{i_1}\neq x_{i_2}\neq\dots\neq x_{i_k}$ and $x_{i_j}\in\{x_1,\dots,x_n\}$, $j=1,\dots,k$, and measure 0 to the rest of X^k. $M_{(k)}$ is the mean distribution of $N_{(k)}$ and hence a measure on $\mathbf{B}^k$, finite if and only if $m_{(k)}<\infty$, σ-finite if and only if $M_{(k)}(A_i^{(k)})<\infty$ for each $A_i^{(k)}$ of some countable measurable partition of X^k. It is then easily seen that

$$M_{(k)}(A^{(k)})=\sum_{n=k}^{\infty}\int_{X^n}\sum_{i_1\neq\dots\neq i_k}\delta(A^{(k)}\,|\,x_{i_1},\dots,x_{i_k})\,P_s^{(n)}(dx^n)$$

$$=\sum_{n=k}^{\infty}\frac{n!}{(n-k)!}P_S^{(n)}(A^{(k)}\times X^{n-k}). \tag{3.12}$$

In particular $M_{(k)}(A^k)$ is the usual kth factorial moment of the random variable $N(A)$. Moment distributions can be expressed in terms of factorial moment distributions; for example,

$$M_2(A^{(2)})=M_{(2)}(A^{(2)})+M(DA^{(2)}), \tag{3.13}$$

where $DA^{(2)}$ is the set of all $x\in X$ such that $x=y$ and $(x,y)\in A^{(2)}$.

In the case of a compound process, where $P^{(n)} = p_n Q^{*n}$, (3.12) takes the simple form

$$M_{(k)}(A^{(k)}) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} p_n Q^{*k}(A^{(k)}) = m_{(k)} Q^{*k}(A^{(k)}); \tag{3.14}$$

hence $M_{(k)}$ exists if and only if $m_{(k)} < \infty$, in which case it is of course finite; the σ-finite case is thus excluded for compound processes. For the Poisson process $M_{(k)} = m^k Q^{*k}$, and for the geometric process $M_{(k)} = k!\, m^k Q^{*k}$ with $m = q/(1-q)$. An example of a population process that has moment distributions of all order which are σ-finite rather than finite is the following: suppose $P^{(n)} = p_n Q_n^{*n}$, where Q_n is a probability distribution on $\mathbf{B}$, $n = 1, 2, \ldots$; then the kth factorial moment distribution

$$M_{(k)} = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} p_n Q_n^{*k}$$

may be defined as a σ-finite distribution even when $m_{(k)}$ is infinite: it is sufficient that there exist a non-decreasing sequence of measurable sets $\{X_i\}$ such that $X_i \uparrow X$, and a non-decreasing sequence of positive numbers $\{a_i\}$ such that $a_i \to 1$ and $Q_n(X_i) \leqslant a_i^n$ for $i = 1, 2, \ldots$, and $n = 1, 2, \ldots$. Thus let $X = [0, 1]$, $g(\lambda) = \sum_{n=0}^{\infty} p_n \lambda^n$, and let Q_n be defined by the cumulative distribution x^n on $[0, 1]$; then

$$M(x) = \int_0^x M(dy) = \sum_{n=1}^{\infty} n p_n x^n = x \frac{dg(x)}{dx},$$

$$M_{(k)}(x_1, \ldots, x_k) = \int_0^{x_1} \int_0^{x_k} M_{(k)}(dy^k) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} p_n x_1^n \ldots x_k^n = x_1^k \ldots x_k^k \left[\frac{\partial^k g(\lambda)}{\partial \lambda^k}\right]_{\lambda = x_1 \ldots x_k}$$

$$(k = 2, 3, \ldots).$$

Take
$$g(\lambda) = a\lambda + (1-a) \sum_{n=2}^{\infty} \frac{\lambda^n}{n(n-1)},$$

where $0 \leqslant a < 1$; then

$$M(x) = ax - (1-a)\, x \log(1-x) \to \infty \quad \text{as} \quad x \to 1,$$

$$M_{(k)}(x_1, \ldots, x_k) = (1-a)\,(k-2)!\, x_1^k \ldots x_k^k (1 - x_1 \ldots x_k)^{1-k} \to \infty \quad \text{as} \quad x_1 \ldots x_k \to 1, \quad (k = 2, 3, \ldots).$$

We have thus constructed a process where $m_{(k)}$ is infinite and yet $M_{(k)}$ exists as a σ-finite distribution for all $k \geqslant 1$.

Let π be a measure on $\mathbf{B}$, let π_n be the nth product measure generated by π on $\mathbf{B}^n$, and suppose that $P_S^{(n)}$ is absolutely continuous with respect to π_n with density f_n; then $M_{(k)}$, if it exists, is absolutely continuous with respect to π_k and its density is

$$\mu_{(k)}(x^k) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \int_{X^{n-k}} f_n(x_1, \ldots, x_n)\, \pi_{n-k}(dx_{k-1} \ldots dx_n). \tag{3.15}$$

The factorial moment density $\mu_{(k)}$ is also called in the literature the *product moment density* (cf. Bartlett [2], p. 79, Ramakrishnan [14]). Note that the moment distributions M_k are not absolutely continuous for $k > 1$, but exhibit "mass" concentrations on subsets of X^k: thus it will be seen from (3.13) that M_2 has such a concentration on the "diagonal" of X^2.

4. Generating Functionals (1)

Let $(\mathfrak{X}, \mathbf{B})$ be the point process measure space generated by the measure space $(X, \mathbf{B})$; let $\mathfrak{M}$ be the space of all bounded measurable complex-valued functions ξ on X; $\mathfrak{M}$ is a linear vector space and becomes a Banach space under the norm

$$\|\xi\| = \sup_{x \in X} |\xi(x)|. \tag{4.1}$$

Consider now the symmetric measurable function $w(x^n) = \xi(x_1) \ldots \xi(x_n)$ on $\mathfrak{X}$. The *probability generating functional* (briefly, p.g.fl.) of a probability distribution P on $\mathbf{B}$ is

$$G[\xi] = Ew = \sum_{n=0}^{\infty} \int_{X^n} \xi(x_1) \ldots \xi(x_n)\, P^{(n)}(dx^n) = \sum_{n=0}^{\infty} G^{(n)}[\xi] = \sum_{n=0}^{\infty} p_n Z^{(n)}[\xi], \tag{4.2}$$

where $p_n = P^{(n)}(X^n)$, $Z^{(n)} = G^{(n)}/p_n$ if $p_n > 0$ and $Z^{(n)} = 0$ if $p_n = 0$. It is clear that:

(1) each $G^{(n)}$ or $Z^{(n)}$ is a functional defined on the whole of $\mathfrak{M}$;

(2) G is a functional defined on a domain $D_G \subset \mathfrak{M}$ which includes the sphere $S_g = \{\xi \mid \|\xi\| \leqslant r_g\}$, where r_g is the radius of convergence of the associated *probability generating function* (briefly p.g.f.) $g(\lambda) = \sum_{n=0}^{\infty} p_n \lambda^n$, which is the value of G at $\xi \equiv \lambda$, where λ is an arbitrary complex variable; hence $r_g \geqslant 1$;

(3) since w is symmetric, two countably equivalent distributions have the same G, and hence G is not altered if we substitute for P in (4.2) its symmetrization P_S. We see thus that to each symmetric probability distribution on $\mathbf{B}$ there corresponds a unique p.g.fl. defined on some domain in $\mathfrak{M}$ which includes the unit sphere $S_0 = \{\xi \mid \|\xi\| \leqslant 1\}$. The functional $Z^{(n)}$ on $\mathfrak{M}$ is the analogue of the nth power of a complex variable in the sense that $Z^{(n)}[\lambda\xi] = \lambda^n Z^{(n)}[\xi]$; in the case of a simple compound process where $P^{(n)} = p_n Q^{*n}$, if

$$Z[\xi] = \int_X \xi(x)\, Q(dx),$$

(1) See Bartlett [1] and [2], Bartlett & Kendall [3] and Kendall [9].

then $Z^{(n)} = Z^n$, and $G[\xi] = g(Z[\xi])$. Thus for a Poisson process $g(\lambda) = \exp m(\lambda - 1)$ and hence

$$G[\xi] = \exp m \int_X [\xi(x) - 1]\, Q(dx) = \exp \int_X [\xi(x) - 1]\, M(dx);$$

for a geometric process $g(\lambda) = [1 - m(\lambda - 1)]^{-1}$ and hence

$$G[\xi] = \left\{1 - m \int_X [\xi(x) - 1]\, Q(dx)\right\}^{-1} = \left\{1 - \int_X [\xi(x) - 1]\, M(dx)\right\}^{-1}.$$

We will now establish the connection between probability generating functionals and functions in the case of counting processes. We *define* the integral of a measurable function θ on X with respect to N to be the random variable

$$\int_X \theta(x)\, N(dx \mid x^{\{n\}}) = \sum_{i=1}^{n} \int_X \theta(x)\, \delta(dx \mid x_i) = \sum_{i=1}^{n} \theta(x_i). \tag{4.3}$$

Hence, taking $\log \xi$ to be the principal branch of the logarithm of ξ,

$$G[\xi] = E\xi(x_1) \ldots \xi(x_n) = E \exp \sum_{i=1}^{n} \log \xi(x_i) = E \exp \int_X \log \xi(x)\, N(dx). \tag{4.4}$$

Let $\{A_1, \ldots, A_k\}$ be a finite measurable partition of X. We define the multivariate p.g.f.

$$g(\lambda_1, A_1; \ldots; \lambda_k, A_k) = E \prod_{i=1}^{k} \lambda_i^{N(A_i)} = \sum_{n_1=0}^{\infty} \ldots \sum_{n_k=0}^{\infty} P_N(N(A_i) = n_i;\ i = 1, \ldots, k)\, \lambda_1^{n_1} \ldots \lambda_k^{n_k}, \tag{4.5}$$

where $\lambda_1, \ldots \lambda_k$ are complex variables of modulus $\leqslant 1$. Now set

$$\xi(x) = \sum_{i=1}^{k} \lambda_i\, \delta(A_i \mid x) \tag{4.6}$$

in (4.4) and (4.2) respectively, and we find

$$G[\xi] = E \exp \sum_{i=1}^{k} \int_{A_i} \log \sum_{j=1}^{k} \lambda_j \delta(A_j \mid x)\, N(dx) = E \prod_{i=1}^{k} \lambda^{N(A_i)} = g(\lambda_1, A_1; \ldots; \lambda_k, A_k) \tag{4.7}$$

$$\begin{aligned} G[\xi] &= \sum_{n=0}^{\infty} \int_{X^n} \prod_{i=1}^{n} \sum_{j=1}^{k} \lambda_j \delta(A_j \mid x_i)\, P_S^{(n)}(dx^n) \\ &= \sum_{n_1=0}^{\infty} \ldots \sum_{n_k=0}^{\infty} \frac{(n_1 + \ldots + n_k)!}{n_1! \ldots n_k!} \lambda_1^{n_1} \ldots \lambda_k^{n_k}\, P_S^{(n_1 + \ldots + n_k)}(A_1^{n_1} \times \ldots \times A_k^{n_k}). \end{aligned} \tag{4.8}$$

By identifying the coefficients of $\lambda_1^{n_1} \ldots \lambda_k^{n_k}$ in (4.7) and (4.8) we find again expression (3.5) for $P_N(N(A_i) = n_i;\ i = 1, \ldots, k)$.

The analogy between probability generating functional and function carries to the generation of probabilities by differentiation: thus for the multivariate p.g.fl. $g(\lambda_1, \dots \lambda_k) = \sum p_{n_1} \dots n_k \lambda_1^{n_1} \dots \lambda_k^{n_k}$, we have

$$n_1! \dots n_k! \, p_{n_1 \dots n_k} = \left[\frac{\partial^{n_1 + \dots + n_k}}{\partial \lambda_1^{n_1} \dots \partial \lambda_k^{n_k}} g(\lambda_1, \dots, \lambda_k) \right]_{\lambda_1 = \dots = \lambda_k = 0}. \tag{4.9}$$

In the case of the p.g.fl. $G[\xi]$, let η, ξ be fixed elements of $\mathfrak{M}$, assume that $\|\eta\| < 1$, and let r be the largest real number such that $\eta + \lambda\xi \in S_g$ for $|\lambda| < r$. Considered as a function on the complex plane Λ

$$G[\eta + \lambda\xi] = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \binom{n}{k} \lambda^k \int_{X^n} \xi(x_1) \dots \xi(x_k) \, \eta(x_{k-1}) \dots \eta(x_n) \, P_S^{(n)}(dx^n)$$

is the limit of a uniformly convergent sequence of polynomials in λ in any closed region interior to the circle $|\lambda| < r$, and hence an analytic function of λ in this circle. We can therefore rearrange the series in the form

$$G[\eta + \lambda\xi] = \sum_{k=0}^{\infty} \lambda^k \sum_{n=k}^{\infty} \binom{n}{k} \int_{X^n} \xi(x_1) \dots \xi(x_k) \, \eta(x_{k+1}) \dots \eta(x_n) \, P_S^{(n)}(dx^n). \tag{4.10}$$

It follows that $G[\eta + \sum_{i=1}^{k} \lambda_i \xi_i]$, where $\|\eta\| < 1$ and $\xi_i \in \mathfrak{M}$, $i = 1, \dots k$, has first partial derivatives with respect to the λ_i in some open region, say D_0, of Λ^k containing the origin, and is therefore an analytic function of the k complex variables $\lambda_1, \dots, \lambda_k$ in D_0, with partial derivatives of all order which are independent of the order of differentiation. This enables us to define the *k-th order variation* of G as follows (cf. Hille and Phillips [8], p. 109)

$$\begin{aligned} \delta^k_{\xi_1 \dots \xi_k} G[\eta] &= \left\{ \frac{\partial^k}{\partial \lambda_1 \dots \partial \lambda_k} G\left[\eta + \sum_{i=1}^{k} \lambda_i \xi_i \right] \right\}_{\lambda_1 = \lambda_2 = \dots = \lambda_k = 0} \\ &= \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \int_{X^n} \xi_1(x_1) \dots \xi_k(x_k) \, \eta(x_{k+1}) \dots \eta(x_n) \, P_S^{(n)}(dx^n). \end{aligned} \tag{4.11}$$

The analogue of (4.9) is then

$$\delta^k_{\xi_1 \dots \xi_k} G[0] = k! \int_{X^k} \xi_1(x_1) \dots \xi_k(x_k) \, P_S^{(k)}(dx^k). \tag{4.12}$$

Set in (4.12) $$\xi_i(x) = \delta(A_i \,|\, x) \quad (i = 1, \dots, k), \tag{4.13}$$

where $A_1, \dots, A_k$ are arbitrary measurable sets in X; then

$$P_S^{(k)}(A_1 \times \dots \times A_k) = \frac{1}{k!} \delta^k_{\xi_1 \dots \xi_k} G[0]. \tag{4.14}$$

This last result shows that G determines P_S uniquely. For suppose P'_S has the same p.g.fl. G. Then by (4.14) P'_S agrees with P_S for all measurable product sets $A_1 \times \dots \times A_k$, and hence on the smallest field containing all such sets. But $\mathbf{B}$ is the minimal σ-field containing this field, and hence by the uniqueness of the extension of a measure, P'_S agrees with P_S on $\mathbf{B}$. We have thus proved:

THEOREM 4.1. *Let* $(\mathfrak{X}, \mathbf{B})$ *be the point process measure space generated by the measure space* $(X, \mathbf{B})$ *and let* $\mathfrak{M}$ *be the space of all bounded measurable complex-valued functions on* X. *There is a one-to-one correspondence given by* (4.2) *between the class of all symmetric probability distributions on* $\mathbf{B}$ *and the class of all probability generating functionals on the unit sphere* S_0 *in* $\mathfrak{M}$.

Carrying the analogy between probability generating functions and functionals still further, we can use the latter to generate factorial moment distributions when they exist. For suppose that $M_{(k)}$ exists as a σ-finite distribution on $\mathbf{B}^k$; if $\xi_1(x_1) \dots \xi_k(x_k)$ is integrable with respect to $M_{(k)}$, then

$$\int_{X^k} \xi_1(x_1) \dots \xi_k(x_k) M_{(k)}(dx^k) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} \int_{X^n} \xi_1(x_1) \dots \xi_k(x_k) P_S^{(n)}(dx^k)$$
$$= \lim_{\eta \to 1} \delta^k_{\xi_1 \dots \xi_k} G[\eta]. \tag{4.15}$$

The second expression in (4.15) follows from the first by (3.12); we can invert summation and integration here because $M_{(k)}$ is the sum of a convergent series of measures. The third expression follows from the second and (4.11) by dominated convergence. Set ξ_i as in (4.13), with $A_1, \dots, A_k$ such that $M_{(k)}(A_1 \times \dots \times A_k) < \infty$ and we find that

$$M_{(k)}(A_1 \times \dots \times A_k) = \lim_{\eta \to 1} \delta^k_{\xi_1 \dots \xi_k} G[\eta]. \tag{4.16}$$

It then follows by the same arguments as where used in the proof of Theorem 4.1 that:

LEMMA 4.2. *The p.g.fl. of a point process uniquely generates all its existing finite or* σ*-finite factorial moment distributions.*

If $r_g > 1$ then all the $M_{(k)}$ are finite and we can set $\eta \equiv 1$ in (4.10) to obtain the expansion of G in terms of factorial moment distributions:

$$G[1+\lambda\xi]=\sum_{k=0}^{\infty}\lambda^k\sum_{n=k}^{\infty}\binom{n}{k}\int_{X^n}\xi(x_1)\ldots\xi(x_k)P_S^{(n)}(dx^n)$$

$$=1+\sum_{k=1}^{\infty}\frac{\lambda^k}{k!}\int_{X^k}\xi(x_1)\ldots\xi(x_k)M_{(k)}(dx^k). \tag{4.17}$$

It is legitimate in this case to substitute $\lambda\xi-1$ for $\lambda\xi$ in the right-hand side of (4.17), and this leads to the relation

$$G[\lambda\xi]=\sum_{n=0}^{\infty}\lambda^n\sum_{k=n}^{\infty}\frac{(-1)^{k+n}}{n!\,(k-n)!}\int_{X^k}\xi(x_1)\ldots\xi(x_n)M_k(dx^k)$$

$$=\sum_{n=0}^{\infty}\lambda^n\int_{X^n}\xi(x_1)\ldots\xi(x_n)P_S^{(n)}(dx^n), \tag{4.18}$$

which yields an expression for $P_S^{(n)}$ in terms of $M_{(k)}$ that is the inverse of expression (3.12) for $M_{(k)}$ in terms of $P_S^{(n)}$, namely:

$$P_S^{(n)}(A^{(n)})=\sum_{k=n}^{\infty}\frac{(-1)^{k+n}}{n!\,(k-n)!}M_{(k)}(A^{(n)}\times X^{k-n}). \tag{4.19}$$

The *characteristic functional* $\varphi[\theta]$ of a point process may be obtained from its p.g.fl. $G[\xi]$ by the substitution $\xi=e^{i\theta}$ where θ is real; thus

$$\varphi[\theta]=E\exp\left(i\int_X\theta(x)N(dx)\right)=\sum_{n=0}^{\infty}\int_{X^n}\exp\left(i[\theta(x_1)+\ldots+\theta(x_n)]\right)P_S^{(n)}(dx^n) \tag{4.20}$$

is a functional on the whole of the space $\mathfrak{M}_R$ of all real-valued bounded measurable functions on X. We can use it to generate the moment distributions when they exist. Let $J[\theta]=\int_X\theta(x)N(dx)$. If M exists as a σ-finite distribution on $\mathbf{B}$ and if θ is integrable with respect to M, then

$$EJ[\theta]=\sum_{n=1}^{\infty}\int_{X^n}\sum_{i=1}^{n}\theta(x_i)P_S^{(n)}(dx^n)=\int_X\theta(x)M(dx). \tag{4.21}$$

If M_k exists as a σ-finite distribution on $\mathbf{B}^k$ and if $\theta_1\ldots\theta_k$ is integrable with respect to M_k, then it follows immediately from (4.21), since $M_k=EN_k$ just as $M=EN$, that

$$E\prod_{i=1}^{k}J[\theta_i]=E\prod_{i=1}^{k}\int_X\theta_i(x)N(dx)=E\int_{X^k}\theta_1(x_1)\ldots\theta_k(x_k)N_k(dx^k)$$

$$=\int_{X^k}\theta_1(x_1)\ldots\theta_k(x_k)M_k(dx^k). \tag{4.22}$$

Hence, since for $\lambda_1,\ldots,\lambda_k$ real

$$\varphi\left[\sum_{j=1}^{k} \lambda_j \theta_j\right] = E \exp i \sum_{j=1}^{k} \lambda_j J[\theta_j] \tag{4.23}$$

is the characteristic function of the k random variables $J[\theta_1], \ldots, J[\theta_k]$, we have

$$\delta^k_{\theta_1 \ldots \theta_k} \varphi[0] = \left\{\frac{\partial^k}{\partial \lambda_1 \ldots \partial \lambda_k} \varphi\left[\sum_{j=1}^{k} \lambda_j \theta_j\right]\right\}_{\lambda_1 = \ldots = \lambda_k = 0}$$

$$= i^k \int_{X^k} \theta_1(x_1) \ldots \theta_k(x_k) M_k(dx^k). \tag{4.24}$$

It follows as before that φ generates uniquely all existing moment distributions. If $m_k < \infty$ for all k and $\sum_{k=0}^{\infty} (i\lambda)^k m_k/k!$ converges, then it is easily seen that

$$\varphi[\theta] = 1 + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int_{X^n} \theta(x_1) \ldots \theta(x_n) M_n(dx^n). \tag{4.25}$$

5. Stochastic Population Processes

We now turn our attention to stochastic processes where the "dependent variable" is a population variable of the kind defined in §§ 2 and 3. Let $(\Omega, \mathbf{B}_\Omega, P_\Omega)$ be a given probability space, let $(X_t, \mathrm{B}_t;\ t \in T)$ be an indexed family of individual measure spaces and $(\mathfrak{X}_t, \mathbf{B}_t;\ t \in T)$ the associated family of population measure spaces. Suppose that for each $t \in T$ there is defined a random variable $x^n(t, \omega)$ on Ω taking its values in $\mathfrak{X}_t$. Let $\mathfrak{X}_T = \prod_{t \in T} \mathfrak{X}_t$ and let $\mathbf{B}_T = \prod_{t \in T} \mathbf{B}_t$; i.e., $\mathbf{B}_T$ is the σ-field generated by the field of all measurable cylinders $A_K \times \mathfrak{X}_{T-K}$, where K is a finite subset of T and $A_K \in \mathbf{B}_K = \prod_{t \in K} \mathbf{B}_t$. Then the measurable transformation $\omega \to x^n(t, \omega)$ yields a probability distribution P_T on $\mathbf{B}_T$, and we may call $(\mathfrak{X}_T, \mathbf{B}_T, P_T)$ a *stochastic population process* in the point process formulation of § 2. For each $t \in T$ and $\omega \in \Omega$, the transformation (3.1) defines a counting measure $N(\cdot\,; \omega)$ on B_t. Let $(\mathfrak{N}_t \mathbf{B}_{N_t};\ t \in T)$ be the family of counting process measure spaces generated in this way; then the transformation yields the stochastic population process $(\mathfrak{N}_T, \mathbf{B}_{N_T}, P_{N_T})$ in the counting process formulation of § 3, where $\mathfrak{N}_T = \prod_{t \in T} \mathfrak{N}_t$ and $\mathbf{B}_{N_T} = \prod_{t \in T} \mathbf{B}_{N_t}$. The standard methods and results of the theory of stochastic processes apply to this case and will be discussed here only briefly (cf. Kolmogorov [10], Doob [5]).

If T is finite, we have a *multivariate population process* and the results of the previous sections extend in an obvious way. Formally, if $T = (1, \ldots, k)$, this extension may be achieved by substituting the vector $n_T = \{n_1, \ldots, n_k\}$ for n and letting $\sum_{n_T}$ stand for $\sum_{n_1=0}^{\infty} \ldots \sum_{n_k=0}^{\infty}$. Thus let $X^{n_T} = \prod_{i=1}^{k} X_i^{n_i}$; then $\mathfrak{X}_T = \sum_{n_T} X^{n_T}$ and if A is a subset of $\mathfrak{X}_T$ then $A = \sum_{n_T} A^{(n_T)}$, where $A^{(n_T)} = A \cap X^{n_T}$. Similarly, let $\mathbf{B}^{n_T} = \prod_{i=1}^k \mathbf{B}_i^{n_i}$; then $\boldsymbol{B}_T$ is the σ-field consisting of all $A = \sum_{n_T} A^{(n_T)}$ such that $A^{(n_T)} \in \mathbf{B}^{n_T}$. A sequence of measures $P^{(n_T)}$ on $\mathbf{B}^{n_T}$ such that $\sum_{n_T} P^{(n_T)}(X^{n_T}) = 1$ determines a unique probability distribution P_T on $\boldsymbol{B}_T$ such that $P_T(A) = \sum_{n_T} P^{(n_T)}(A^{(n_T)})$ for every $A = \sum_{n_T} A^{(n_T)}$ belonging to $\boldsymbol{B}_T$. Let N_i be the associated counting measure on $\mathbf{B}_i$, let $N_{(n_i)}$ be defined in terms of N_i by (3.11) and form the product measure $N_{(n_T)} = N_{(n_1)} \times \ldots \times N_{(n_k)}$; then the generalization of expression (3.12) for factorial moment distributions is

$$M_{(n_T)}(A^{(n_T)}) = EN_{(n_T)}(A^{(n_T)}) = \sum_{r_T} \prod_{i=1}^{k} \frac{(n_i + r_i)!}{n_i!} P^{(n_T + r_T)}(A^{(n_T)} \times X^{r_T}). \tag{5.1}$$

Let ξ_i be a bounded measurable function on X_i, and let $\xi_T = \{\xi_1, \ldots, \xi_k\}$. Then the p.g.fl. of P_T is by definition

$$G[\xi_T] = \sum_{n_T} \int_{X^{n_T}} \prod_{i=1}^{k} \xi_i(x_1^{(i)}) \ldots \xi_i(x_{n_i}^{(i)}) P^{(n_T)}(dx^{n_T}), \tag{5.2}$$

where $x^{n_T} = \{x^{n_1}, \ldots, x^{n_k}\}$ stands for a point in $\mathfrak{X}_T$, $x^{n_i} = \{x_1^{(i)}, \ldots, x_{n_i}^{(i)}\}$ for a point in $X_i^{n_i}$ and $x^{(i)}$ for a point in X_i.

The concepts and properties of stochastic independence, conditional distributions and conditional expectations all carry over to population processes. Thus if $\mathbf{x}, \mathbf{y}$ are two population processes with individual measure spaces respectively $(X, \mathbf{B}_x)$, $(Y, \mathbf{B}_y)$ and associated population measure space respectively $(\mathfrak{X}, \boldsymbol{B}_x)$, $(\mathfrak{Y}, \boldsymbol{B}_y)$, then a probability distribution for the $\mathbf{y}$ population conditional on the state of the $\mathbf{x}$ population is a function Q on $\boldsymbol{B}_y \times \mathfrak{X}$ such that for every fixed $x^k \in \mathfrak{X}$, $Q(\cdot \mid x^k)$ is a probability distribution on $\boldsymbol{B}_y$ and for every fixed $A \in \boldsymbol{B}_y$ $Q(A \mid \cdot)$ is a measurable function on $\mathfrak{X}$. The corresponding conditional p.g.fl. is

$$G[\xi \mid x^k] = \sum_{n=0}^{\infty} \int_{Y^n} \xi(y_1) \ldots \xi(y_n) Q^{(n)}(dy^n \mid x^k). \tag{5.3}$$

Of particular importance in the theory of cluster and multiplicative population processes (see § 7) is the case where Q has the *multiplicative property*; this may be expressed as follows: let $N(A \mid x^k)$ be the value at $A \in \mathbf{B}_y$ of the counting process associated to the point process with distribution $Q(\cdot \mid x^k)$; then for all $A \in \mathbf{B}_y$

$$N(A \mid x^k) = \sum_{i=1}^{k} N(A \mid x_i),$$

where $N(A \mid x_1), \ldots, N(A \mid x_k)$ are mutually independent random variables. If this is the case, then it is easily seen that the conditional p.g.fl. (5.3) factorizes as follows:

$$G[\xi \mid x^k] = \prod_{i=1}^{k} G[\xi \mid x_i]. \tag{5.4}$$

Returning to the case where T is infinite, if K is a finite subset of T, then we obtain by projection the multivariate process $(\mathfrak{X}_K, \mathbf{B}_K, P_K)$, where $P_K(A) = P_T(A \times \mathfrak{X}_{T-K})$ for every $A \in \mathbf{B}_K$. The family of all such multivariate processes for all finite subsets of T is *consistent*, in the sense that if $K' \subset K$, then $P_{K'}(A) = P_K(A \times \mathfrak{X}_{K-K'})$ for all $A \in \mathbf{B}_{K'}$. A stochastic population process is often specified by means of a consistent family of multivariate distributions $\{P_K\}$ defined for all finite subsets K of T; hence we need an extension theorem which will guarantee that such a family has a unique extension to a probability distribution P_T on $\mathbf{B}_T$. Kolmogorov's extension theorem is proved (Kolmogorov [10], p. 29) for stochastic processes which take their value in a Euclidean space. It is not difficult to see that this proof can be generalized to the present case if each X_t is a Euclidean space with the usual topology, or even more generally a locally compact Hausdorff space, and each $\mathbf{B}_t$ is the σ-field generated by the Borel sets of X_t. We have in fact available a further generalization due to I. E. Segal [15] which as applied to the present case states that given a consistent family $\{P_K\}$, and without any restrictions on the X_t, there exists a probability space $(\Omega, \mathbf{B}_\Omega, P_\Omega)$ and an indexed family of "generalized" random variables on this space with index set T whose joint probability distributions for finite sets of indices agree with the family $\{P_K\}$.

6. σ-finite Population Processes

So far we have only considered populations whose total size is finite with probability one; we will now extend the theory to the case of populations that can be infinite with positive probability. This is most conveniently done in terms of counting processes: for a given measure space $(X, \mathbf{B})$, let $\mathcal{N}_\omega$ be a space of *uniformly σ-finite counting measures* on $\mathbf{B}$: by this we mean the set of all integral-valued completely additive measures N_ω on $\mathbf{B}$ such that $N_\omega(X_k) < \infty$ for each X_k of a given *fixed* non-decreasing sequence of measurable sets $\{X_i\}$ such that $X_i \uparrow X$. A *σ-finite counting process* is by definition a probability space $\mathrm{N}_\omega = (\mathcal{N}_\omega, \mathbf{B}_\omega, P_\omega)$, where $\mathbf{B}_\omega$ is a suitably defined σ-field of sets in $\mathcal{N}_\omega$ and P_ω is a probability distribution on $\mathbf{B}_\omega$. Given such a probability space, let $\mathcal{N}_k$ be the restriction of $\mathcal{N}_\omega$ to X_k: i.e., $\mathcal{N}_k$ is the set of all restrictions of elements of $\mathcal{N}_\omega$ to X_k, and is therefore the space of all *finite* counting measures on the σ-field $\mathbf{B}_k$ of all measurable subsets of X_k; let $\boldsymbol{B}_k$ be the restriction of $\boldsymbol{B}_\omega$ to X_k: i.e., the class of all sets S_k in $\mathcal{N}_k$ whose continuation S_k^ω to X belongs to $\boldsymbol{B}_\omega$; finally, let Q_k be the function on $\boldsymbol{B}_k$ whose value at S_k is $Q_k(S_k) = P_\omega(S_k^\omega)$. Clearly $\boldsymbol{B}_k$ is a σ-field of sets in $\mathcal{N}_k$, Q_k is a probability distribution on $\boldsymbol{B}_k$ and $\mathbf{N}_k = (\mathcal{N}_k, \boldsymbol{B}_k, Q_k)$ is a *finite* counting process, which we may term the X_k-restriction of $\mathbf{N}_\omega$. The sequence $\{\mathbf{N}_i\}$ (or equivalently the sequence $\{Q_i\}$) is *consistent under restrictions* in the sense that if $j > k$ and S_k^j is the continuation of the set $S_k \in \boldsymbol{B}_k$ to X_j, then $S_k^j \in \mathbf{B}_j$ and $Q_k(S_k) = Q_j(S_k^j)$. Conversely, let $(\mathcal{N}_k, \boldsymbol{B}_k)$ be the finite counting process measure space generated as in § 3.1 by $(X_k, \boldsymbol{B}_k)$ and let Q_k be a probability distribution on $\boldsymbol{B}_k$. If the sequence $\{Q_i\}$ is consistent under restrictions, then we can apply Kolmogorov's extension theorem. The measurable cylinders of § 5 are here the *measurable continuations* S_k^ω, where $S_k \in \boldsymbol{B}_k$. The class $\boldsymbol{C}_\omega$ of all such measurable continuations is obviously a field of sets in $\mathcal{N}_\omega$; we define a function Q_ω on $\boldsymbol{C}_\omega$ as follows; if $S \in \boldsymbol{C}_\omega$, then there exists an integer k such that $S = S_k^\omega$, where $S_k \in \boldsymbol{B}_k$; we set $Q_\omega(S) = Q_k(S_k)$. This definition is consistent: for if $S = S_k^\omega = S_j^\omega (j > k)$, then clearly $S_j = S_k^j$ and hence because of the consistency condition $Q_k(S_k) = Q_j(S_j)$. Clearly Q_ω is finitely additive and normalized to unity. If we now take $\boldsymbol{B}_\omega$ to be the σ-field generated by $\boldsymbol{C}_\omega$, then by Kolmogorov's theorem Q_ω has a unique extension to a probability distribution P_ω on $\boldsymbol{B}_\omega$. We have thus proved:

THEOREM 6.1. *A sequence of finite counting processes $\{\mathbf{N}_k\}$ consistent under restrictions determines a unique σ-finite counting process $\mathbf{N}_\omega$ such that $\mathbf{N}_k$ is the X_k-restriction of $\mathbf{N}_\omega$, $k = 1, 2 \ldots$. Conversely if $\mathbf{N}_\omega$ is a σ-finite counting process with finite X_k-restriction $\mathbf{N}_k$, then the sequence $\{\mathbf{N}_k\}$ is consistent under restrictions.*

Using the same notation as in the last paragraph, we *define* the p.g.fl. of the σ-finite counting process $\mathbf{N}_\omega$ to be

$$G_\omega[\xi] = E \exp \int_X \log \xi(x) N_\omega(dx). \tag{6.1}$$

This functional is defined on a certain domain D_ω in the space $\mathcal{M}$ of all bounded measurable complex-valued functions ξ on X. The p.g.fl.

$$G_k[\xi] = E \exp \int_{X_k} \log \xi(x) N_k(dx) \tag{6.2}$$

of the X_k-restriction $\mathbf{N}_k$ of $\mathbf{N}_\omega$ is obtained from G_ω by setting $\xi(x) = 1$ for all $x \in X - X_k$ in (6.1). The sequence $\{G_i\}$ defined in this way is consistent under restrictions in the sense that whenever $j > k$, G_k agrees with G_j if $\xi(x) = 1$ for all $x \in X_j - X_k$. It can now be seen to follow from Theorems 4.1 and 6.1 that:

THEOREM 6.2. *A sequence of finite counting processes $\{\mathbf{N}_i\}$ is consistent under restrictions if and only if the corresponding sequence of p.g.fls. $\{G_i\}$ is so. A functional G_ω on a domain D_ω in $\mathfrak{M}$ is the p.g.fl. of a σ-finite counting process $\mathbf{N}_\omega$ if and only if there exists a sequence of measurable sets $\{X_i\}$ with $X_i \uparrow X$ such that:* (1) *D_ω contains at least all ξ such that $\|\xi\| \leqslant 1$ and, for some k, $\xi(x) = 1$ for all $x \in X - X_k$;* (2) *each X_k-restriction G_k of G_ω is the p.g.fl. of a finite counting process and the sequence $\{G_i\}$ is consistent under restrictions. If this is true then the corresponding sequence of finite counting processes $\{\mathbf{N}_i\}$ determines $\mathbf{N}_\omega$ uniquely.*

Two examples which will serve to illustrate the theory are the Poisson and geometric σ-finite population processes, with p.g.fls. respectively

$$G_\omega[\xi] = \exp \int_X [\xi(x) - 1] M(dx), \tag{6.3}$$

$$G_\omega\{\xi\} = \left\{1 - \int_X [\xi(x) - 1] M(dx)\right\}^{-1}, \tag{6.4}$$

where M is a σ-finite measure on $\mathbf{B}$. It can be seen that both functionals satisfy the conditions of Theorem 6.2; in fact if $M(X_k) < \infty$ for $k = 1, 2, \ldots$ and $X_k \uparrow X$, then the X_k-restrictions form consistent sequences of finite respectively Poisson and geometric population processes.

The definition of moment and factorial moment distributions given in § 3 for finite counting processes extends immediately to the σ-finite case. Thus if the X_k-restrictions $\mathbf{N}_k$ of $\mathbf{N}_\omega$ have finite mean distributions $M^{(k)} = EN_k$, then the mean distribution $M = E N_\omega$ of N_ω is a σ-finite measure on $\mathbf{B}$ such that $M(X_k) = M^{(k)}(X_k) < \infty$, $k = 1, 2, \ldots$; in fact for every $A \in B$, we have

$$M(A) = \lim_{k \to \infty} M^{(k)}(A \cap X_k)$$

in the extended number system. Similar considerations apply to the higher order moment distributions. If follows that the result of § 4 on the generation of moment distributions by generating functionals also extend to the σ-finite case. Take for example the σ-finite Poisson and geometric processes. Their mean distributions are the σ-finite measures M in respectively (6.3) and (6.4), and their kth order factorial moment distributions are respectively $M_{(k)} = M^{*k}$ and $M_{(k)} = k!\, M^{*k}$.

We have so far considered σ-finite population processes as counting processes. Associated with each finite X_k-restriction $\mathbf{N}_k$ of the σ-finite counting process $\mathbf{N}_\omega$ will be a finite symmetric point process, say x_k; we may therefore say that the sequence $\{x_k\}$ is consistent under restrictions, and that it determines a unique σ-finite symmetric point process x_ω associated with $\mathbf{N}_\omega$. The "sample space" of x_ω is the class $\mathfrak{X}_\omega$ of all countable sets of points $\{x_i\}$ in X such that $\{x_i\} \cap X_k$ is finite for $k = 1, 2, \ldots$. The connection between $\mathbf{N}_\omega$ and x_ω is then expressed by:

THEOREM 6.3. *The relation*

$$N_\omega(A \mid \{x_i\}) = \sum_{j=1}^{\infty} \delta(A \mid x_j) \tag{6.5}$$

defines a one-to-one correspondence between $\mathfrak{X}_\omega$ and $\mathfrak{N}_\omega$.

Proof. Clearly if N_ω is defined by (6.5) then it is an element of $\mathfrak{N}_\omega$; the converse part of the theorem is an easy extension of Theorem 3.1.

Let P_k be the probability distribution of the associated point process $\mathbf{x}_k$, $k = 1, 2, \ldots$; then

$$q = \lim_{r\to\infty} \lim_{k\to\infty} \left[1 - \sum_{n=0}^{r} P_k^{(n)}(X_k^n)\right] \tag{6.6}$$

can be interpreted as the probability that the total population is infinite. The Poisson and geometric processes for example have $q=0$ if M is finite and $q=1$ if $M(X)=\infty$. A trivial example of a process with q positive but less than one is a mixture with p.g.fl. $G = aG_1 + (1-a)\ G_2\ (0<a<1)$, where G_1 is the p.g.fl. of a σ-finite process with $q=1$ and G_2 that of a finite process.

If X_ν is any measurable subset of X, we can define the X_ν-restriction $\mathbf{N}_\nu$ of $\mathbf{N}_\omega$ exactly as before: we call X_ν **N**-*finite* if N_ν is a finite counting process. Clearly the joint probability distribution of any finite collection $\mathbf{N}_1, \ldots, \mathbf{N}_k$ of finite restrictions of $\mathbf{N}_\omega$ is uniquely determined by the distribution of $\mathbf{N}_\omega$. In terms of p.g.fls., we can generalize expression (6.2) as follows: let ξ_ν be a bounded measurable function on X_ν and let ξ'_ν on X be such that $\xi'_\nu(x) = \xi_\nu(x)$ for all $x \in X_\nu$ and $\xi'_\nu(x) = 1$ for all $x \in X - X_\nu$, $\nu = 1, \ldots, k$; then the p.g.fl. of the joint distribution of the $\mathbf{N}_\nu$ is

$$G[\xi_1, \ldots, \xi_k] = G_\omega\left[\sum_{\nu=1}^{k} \xi'_\nu\right]. \tag{6.7}$$

Suppose now that $X_1, \ldots, X_k$ are mutually disjoint. Let $X_K = \sum_{\nu=1}^{k} X_\nu$, let $\mathbf{N}_K$ be the X_K-restriction of $\mathbf{N}_\omega$ and let P_K be the probability distribution of the associated symmetric point process $\mathbf{x}_K$. It is then easily seen that $G[\xi_1, \ldots, \xi_k]$ can be expressed in terms of P_K as follows:

$$G[\xi_1, \dots \xi_k] = \sum_{n_1=0}^{\infty} \dots \sum_{n_k=0}^{\infty} \frac{(n_1 + \dots + n_k)!}{n_1! \dots n_k!} \int_{X_1^{n_1} \times \dots \times X_k^{n_k}} \prod_{j=1}^{k}$$
$$\times \xi_j(x_1^{(j)}) \dots \xi_j(x_n^{(j)}) P^{(n_1 + \dots + n_k)}(dx^{(n_1 + \dots + n_k)}); \qquad (6.8)$$

this is in fact a generalization of expression (4.8) and the proof is similar.

7. Examples

We shall now illustrate the general theory by means of examples of special types of population processes. The treatment will be very brief throughout and we shall omit the proofs. Details of most of the material in this section will be published elsewhere. (1)

1. **Cluster processes.** (2) A *cluster process* is a population process where the individuals are grouped in independent clusters. Each cluster is itself a population whose state is characterized by the ordered pair (x, y^n), where x is characteristic of the cluster as a whole and y^n means that the cluster contains n individuals in states $y_1, \dots, y_n$. Let $(X, \mathbf{B}_x)$ be the measure space of all cluster state variables x, $(\mathcal{Y}, \mathcal{B}_y)$ that of the population state variables y^n, and let $Q(\cdot \mid x)$ be the probability distribution of the population contained in a cluster in state x: Q is a conditional distribution on $\mathcal{B}_y \times X$. The independence of the clusters is expressed by the fact that the distribution of any k clusters in states $x^k = (x_1, \dots, x_k)$ is the conditional probability

$$Q^{*k}(\cdot \mid x^k) = Q(\cdot \mid x_1) \times \dots \times Q(\cdot \mid x_k) \qquad (7.1)$$

on $\mathcal{B}_y^k \times X^k$. Let

$$Z[\xi \mid x] = \sum_{n=0}^{\infty} \int_{Y^n} \xi(y^n, x) Q^{(n)}(dy^n \mid x). \qquad (7.2)$$

and let P be the cluster population distribution relative to the states x, with p.g.fl. G_p; then it is easily seen to follow from (7.1) that the p.g.fl. of the whole cluster process is

$$G[\xi] = \sum_{n=0}^{\infty} \int_{X^n} Z[\xi \mid x_1] \dots Z[\xi \mid x_n] P^{(n)}(dx^n) = G_p\{Z[\xi \mid \cdot]\}. \qquad (7.3)$$

The theory is easily extended to higher order cluster processes, where the individuals

---

(1) *Note added in proof.* See J. E. Moyal, *Proc. Roy. Soc. A*, 266 (1962), 518–526.

(2) See Neyman and Scott (13).

are members of 1st order clusters, which are themselves members of 2nd order clusters, and so on.

One sees from (7.3) that a cluster process is a generalization of a compound process: in fact, if X reduces to a single state, p_n is the probability of n clusters, $g(\lambda)=\sum_{n=0}^{\infty} p_n \lambda^n$, then

$$G[\xi]=\sum_{n=0}^{\infty} p_n \left[\sum_{k=0}^{\infty} \int_{Y^k} \xi(y^k)\, Q^{(k)}(dy^k)\right]^n = g(Z[\xi]); \tag{7.4}$$

i.e., we have here a compound process whose "individuals" are independent populations with distribution Q on $\boldsymbol{B}_y$. If Y reduces to a single state, then $\xi(y^n, x)=\xi^n(x)$; let $q_n(x)$ be the probability of n individuals in a cluster in state x, $g(\xi \mid x)=\sum_{n=0}^{\infty} q_n(x)\, \xi^n(x)$; then

$$G[\xi]=\sum_{n=0}^{\infty} \int_{X^n} g(\xi \mid x_1) \ldots g(\xi \mid x_n)\, P^{(n)}(dx^n)=G_p[g(\xi \mid \cdot)]; \tag{7.5}$$

we call a process of this type a *simple cluster process.*

As illustrations we consider the *Poisson and geometric cluster processes,* with p.g.fls. respectively

$$G[\xi]=\exp \int_X \{Z[\xi] \mid x]-1\}\, M(dx)=\exp \sum_{n=1}^{\infty} \int_{Y^n \times X} [\xi(y^n, x)-1]\, M^{(n)}(dy^n\, dx), \tag{7.6}$$

$$G[\xi]=[1-\int_X \{Z[\xi \mid x]-1\}\, M(dx)]^{-1}=\left[1-\sum_{n=1}^{\infty} \int_{Y^n \times X} [\xi(y^n, x)-1]\, M^{(n)}(dy^n\, dx)\right]^{-1}, \tag{7.7}$$

where $M^{(n)}$ is the measure defined for each $A \in B_y^n \times B_x$ by

$$M^{(n)}(A)=\int_X \int_{Y^n} \delta(A \mid y^n, x)\, Q^{(n)}(dy^n \mid x)\, M(dx), \quad (n=1,2, \ldots). \tag{7.8}$$

The factorization of G in (7.6) shows that a Poisson cluster process may be regarded as the sum of an infinite series of independent Poisson processes with mean distributions $M^{(n)}$. If Y reduces to a single state, we obtain the *simple Poisson cluster process* with p.g.fl.

$$G[\xi]=\exp \sum_{n=1}^{\infty} \int_X [\xi^n(x)-1]\, q_n(x)\, M(dx). \tag{7.9}$$

2. **Counting processes with independent elements.** The simple Poisson cluster process has a feature which is immediately apparent from the form of its p.g.fl. (7.9), namely, that it constitutes a *counting process N with independent elements*: by which

we mean that for any finite measurable partition $\{X_1, \ldots, X_k\}$ of X, the X_i-restrictions N_i of N are mutually independent; for clearly the p.g.fl. of the collection $\{N_1, \ldots, N_k\}$ is

$$G[\xi_1, \ldots, \xi_k] = \exp \sum_{n=0}^{\infty} \int_X \left[\sum_{i=1}^{k} \xi_i^n(x) \delta(X_i \mid x) - 1 \right] q_n(x) M(dx)$$

$$= \exp \sum_{i=1}^{k} \sum_{n=0}^{\infty} \int_{X_i} [\xi_i^n(x) - 1] q_n(x) M(dx) = G[\xi_1] \ldots G[\xi_k], \tag{7.10}$$

a result which is still true if M is a σ-finite measure on $\mathbf{B}$. This is in a sense the most general distribution for a counting process with independent elements: that not all such processes conform to it is evident from the counter example of a counting process where $N(A) = \sum_{x_k \in A} n_k$, where the n_k are a countable collection of mutually independent non-negative integral-valued random variables and each n_k is attached to a point $x_k \in X$. Such trivial cases are excluded in the following lemma:

Let N be a σ-finite counting process with independent elements such that $P_N\{N(A) = 0\} > 0$ for every measurable N-finite set A; then $M_0(A) = -\log P_N\{N(A) = 0\}$ is a σ-finite measure on $\mathbf{B}$. *If M_0 is nonatomic, then the p.g.fl. of N is of the form* (7.9) *with $M_0 = M$.*

Since every singleton $\{x\}$ is an atom of M_0, the condition that M_0 be nonatomic implies that $M_0\{x\} = 0$ and hence that $P_N\{N\{x\} = 0\} = 1$, which excludes the counter example above. Note that if X is the real line, then this theorem is a special case of the Lévy-Kolmogorov decomposition of infinitely divisible distributions (cf. Feller [6], p. 271).

3. **Time-dependent Markov population processes.** A time-dependent Markov population process is characterized by its *transition probability* $P(A, t \mid x^k, s)$ (the probability of a transition from state x^k at time s to some state $x^n \in A$ at time $t \geqslant s$) defined for all $t \geqslant s$ and satisfying the Chapman-Kolmogorov equation

$$P(A, t \mid x^k, s) = \sum_{j=0}^{\infty} \int_{X^j} P(A, t \mid y^j, u) P^{(j)}(dy^j, u \mid x^k, s), \quad (t \geqslant u \geqslant s). \tag{7.11}$$

The "time-axis" T may be the real line (continuous time processes) or the set of all integers (Markov chains). For continuous time processes, transitions which involve a change in the *size* of the population must be of the nature of sudden "jumps", so that the general theory of discontinuous Markov processes (Moyal [12]) is applicable. The transition probability P satisfies the integral equation

$$P(A, t \mid x^k, s) = P_0(A, t \mid x^k, s) + \sum_{j=0}^{\infty} \int_s^t \int_{X^j} P(A, t \mid y^j, u) Q^{(j)}(dy^j, du \mid x^k, s) \tag{7.12}$$

(symbolically, $P = P_0 + P \ast Q$), where P_0 is the probability of a transition $x^k \to A$ in (s, t) *without jumps*, and hence conserving the total number of individuals, Q is the joint probability of the 1st jump time and consequent state (i.e., the state of the population resulting from the jump) conditional on the "initial" state x^k and time s. Given P_0 and Q satisfying the consistency conditions given in Moyal (12), the problem is to solve (7.12) for P. The solution is of the form

$$P = \sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} P_0 \ast Q_n, \tag{7.13}$$

where $\{Q_n\}$ is the Markov chain obtained by iteration of $Q (Q_{n+1} = Q_n \ast Q)$; P_n is the transition probability involving exactly n jumps; Q_n is the probability of the nth jump time and consequent state conditional on some "initial" state and time; $\sigma_n(t \mid x^k, s) = Q_n(\mathfrak{X}, t \mid x^k, s)$ is therefore the cumulative probability distribution of the nth jump time conditional on (x^k, s) and hence the probability of $\leqslant n$ jumps in (s, t) conditional on x^k at s; $\sigma_\infty = \lim_{n\to\infty} \sigma_n$ is interpreted as the probability of infinitely many jumps in (s, t) conditional on x^k at s. The solution (7.13) is "honest" (i.e., P is normalized to unity) and unique if and only if $\sigma_\infty \equiv 0$.

If the process is purely discontinuous with finite "jump" rate $q(x^k, t)$ and probability $W(A \mid x^k, t)$ of a transition $x^k \to A$ given a "jump" at t, then under certain regularity conditions (cf. Moyal (12)) (7.12) is equivalent to the "backward" integro-differential equation

$$\left(-\frac{\partial}{\partial s} + q(x^k, s)\right) P(A, t \mid x^k, s) = q(x^k, s) \sum_{j=0}^{\infty} \int_{X^j} P(A, t \mid y^j, s) W^{(j)}(dy^j \mid x^k, s), \tag{7.14}$$

or in terms of transition generating functionals

$$\left(-\frac{\partial}{\partial s} + q(x^k, s)\right) G[\xi, t \mid x^k, s] = q(x^k, s) \sum_{j=0}^{\infty} \int_{X^j} G[\xi, t \mid y^j, s] W^{(j)}(dy^j \mid x^k, s). \tag{7.15}$$

As an example, we consider the *time dependent Poisson process* with "birth-rate" $q(t)$ and probability $W(A, t)$ of the "created" population at each birth, both independent of the state of the population at t. Then clearly the p.g.fl. of the process must be of the form

$$G[\xi, t \mid x^k, s] = \xi(x_1) \ldots \xi(x_k) G_{(0)}[\xi, t, s], \tag{7.16}$$

where $G_{(0)}$ is the transition p.g.fl. conditional on 0 individuals at s. The "backward equation" becomes

$$-\frac{\partial}{\partial s} G_{(0)} = q(s) G_{(0)} \sum_{j=1}^{\infty} \int_{X^j} [\xi(x_1) \dots \xi(x_j) - 1] W^{(j)}(dy^j, s) = q(s) G_{(0)}[G_W - 1], \tag{7.17}$$

where G_W is the p.g.fl. of W. The solution of (7.17) with the initial condition $G_{(0)}[\xi, s, s] = 1$ is

$$G_{(0)}[\xi, t, s] = \exp \int_s^t \{G_W[\xi, u] - 1\} q(u) du, \tag{7.18}$$

which by comparison with (7.6) is seen to be the p.g.fl. of a cluster Poisson process with mean density of clusters q on (s, t).

4. **Multiplicative population processes.** A multiplicative population process(1) is loosely speaking a Markov process where the individuals at a given time, say s, are the "ancestors" of mutually independent populations at times $t \geqslant s$. More precisely, a multiplicative process is characterized by the fact that its transition probability P has for all $t \geqslant s$ the multiplicative property defined at the end of § 5; hence by (5.4) the transition p.g.fl. G conditional on k "ancestors" in state $x_1, \dots, x_k$ always factorizes as follows:

$$G[\xi, t \mid x^k, s] = \prod_{i=1}^{k} G[\xi, t \mid x_i, s], \quad (t \geqslant s). \tag{7.19}$$

Such a process is therefore uniquely characterized by the transition probability $P(\cdot, t \mid x, s)$ or the p.g.fl. $G[\xi, t \mid x, s]$ conditional on the state of a single "ancestor". If follows from the Chapman-Kolmogorov equation (7.11) that G satisfies the functional relation

$$\begin{aligned} G[\xi, t \mid x, s] &= \sum_{n=0}^{\infty} \int_{X^n} \prod_{i=1}^{n} G[\xi, t \mid y_i, u] P(dy^n, u \mid x, s) \\ &= G[G[\xi, t \mid \cdot, u], u \mid x, s], \quad (t \geqslant u \geqslant s). \end{aligned} \tag{7.20}$$

It can be shown that the mean distribution M of P satisfies an analogue of the Chapman-Kolmogorov equation:

$$M(A, t \mid x, s) = \int_X M(A, t \mid y, u) M(dy, u \mid x, s), \quad (t \geqslant u \geqslant s). \tag{7.21}$$

Similar relations may be found for higher order factorial moment distributions.

(1) See Bartlett [2] and Harris [7], where further references will be found.

In the continuous time case, the integral equation (7.12), expressed in terms of p.g.fls., takes the form

$$G[\xi, t \mid x, s] = G_0[\xi, t \mid x, s] + H\{G[\xi, t\ \cdot\,, \cdot\,], t \mid x, s\}, \tag{7.22}$$

where G_0 is the p.g.fl. of P_0, and H is the p.g.fl. of Q:

$$H[\eta, t \mid x, s] = \sum_{n=0}^{\infty} \int_s^t \int_{X^n} \eta(y_1, u) \ldots \eta(y_u, u)\, Q^{(n)}(dy^n\, du \mid x, s). \tag{7.23}$$

The mean distribution M of P satisfies the integral equation (analogous to (7.22))

$$M(A, t \mid x, s) = M^{(0)}(A, t \mid x, s) + \int_s^t \int_X M(A, t \mid y, u)\, \Lambda(dy\, du \mid x, s), \tag{7.24}$$

(symbolically, $M = M^{(0)} + M * \Lambda$), where Λ is the mean distribution of Q and $M^{(0)} = P_0$. Let $\{\Lambda^{(n)}\}$ be the sequence defined recursively by $\Lambda^{(n+1} = \Lambda^{(n)} * \Lambda$; then the smallest non-negative solution of (7.24) is of the form (corresponding to (7.13))

$$M = \sum_{n=0}^{\infty} M^{(n)} = \sum_{n=0}^{\infty} M^{(0)} * \Lambda^{(n)}. \tag{7.25}$$

Similar results hold for the higher-order factorial moment distributions $M_{(n)}$ of P. For example $M_{(2)}$ satisfies the equation $M_{(2)} = M_{(2)}^{(0)} + M_{(2)} * \Lambda$, where

$$M_{(2)}^{(0)}(A, t \mid x, s) = \int_s^t \int_{X^2} M^{*2}(A, t \mid y^2, u)\, \Lambda_{(2)}(dy^2\, du \mid x, s),$$

$\Lambda_{(2)}$ is the 2nd order factorial moment distribution of Q and

$$M^{*2}(\,\cdot\,, t \mid x_1, x_2, s) = M(\,\cdot\,, t \mid x_1, s) \times M(\,\cdot\,, t \mid x_2, s),$$

whose minimal non-negative solution is

$$M_{(2)} = \sum_{n=0}^{\infty} M_{(2)}^{(n)} = \sum_{n=0}^{\infty} M_{(2)}^{(0)} * \Lambda^{(n)} \tag{7.26}$$

As an example, consider the "birth-and-death" purely discontinuous, time-homogeneous multiplicative population process with constant jump rate q and probability $W(A \mid x)$ of a transition $x \to A$ given a jump. The "backward" equation (7.15) then takes the form

$$\left(\frac{\partial}{\partial t} + q\right) G[\xi \mid x; t] = qK\{G[\xi \mid \cdot\,; t] \mid x\}, \tag{7.27}$$

where K is the p.g.fl. of W. Equations (7.25) and (7.26) then yield

$$M(A\,|\,x;t) = \sum_{n=0}^{\infty} \frac{(qt)^n}{n!} e^{-qt}\,\Gamma^{(n)}(A\,|\,x); \tag{7.28}$$

$$M_{(2)}(A\,|\,x;t) = \int_0^t q\,du \int_X \int_{X^2} M^{*2}(A\,|\,y^2; t-u)\,\Gamma_{(2)}(dy^2\,|\,z)\,M(dz\,|\,x;u), \tag{7.29}$$

where Γ is the mean, $\Gamma_{(2)}$ the 2nd order factorial moment distribution of W, and $\{\Gamma^{(n)}\}$ is defined recursively by $\Gamma^{(n+1)} = \Gamma^{(n)} * \Gamma$.

Consider now the particularly simple case where at each "birth" the "parent" remains in the same state x, and only one "newborn" is produced in a state whose probability distribution Φ is independent of the "parent's" state x. Let λ, μ be the (constant) birth and death rates; then (7.27) becomes

$$\left(\frac{\partial}{\partial t} + \lambda + \mu\right) G[\xi\,|\,x;t] = \mu + \lambda G[\xi\,|\,x;t] \int_X G[\xi\,|\,y;t]\,\Phi(dy), \tag{7.30}$$

whose solution can be expressed in a closed form:

$$G[\xi\,|\,x;t] = \begin{cases} 1 + e^{-\mu t}[\xi - Z + (Z-1)e^{\lambda t}]\left[1 + \dfrac{\lambda}{\mu-\lambda}(Z-1)\left(e^{(\lambda-\mu)t}-1\right)\right]^{-1} & \text{if } \lambda \neq \mu, \\ 1 + (\xi - Z)e^{-\lambda t} + (Z-1)[1 - \lambda t(Z-1)]^{-1} & \text{if } \lambda = \mu, \end{cases} \tag{7.31}$$

where $Z[\xi] = \int_X \xi(x)\,\Phi(dx)$. Let $\mu = 0$ and we obtain the p.g.fl. of a geometric population process with mean distribution $M = (e^{\lambda t} - 1)\,\Phi$:

$$G[\xi\,|\,x;t] = \xi(x)\,[1 - (e^{\lambda t} - 1)\int_X [\xi(x) - 1]\,\Phi(dx)]^{-1}. \tag{7.32}$$

We say that the process is *time-homogeneous* if $P(A, t\,|\,x, s)$ depends only on $t - s$; when this is the case we write $P(\cdot\,|\,\cdot\,; t)$ for the transition probability, $G[\xi\,|\,x;t]$ for its p.g.fl. and $p^{(n)}(x;t) = P^{(n)}(X^n\,|\,x;t)$. One is interested in the asymptotic properties of these distributions as $t \to \infty$. The following results can be proved: (1) For fixed x, the *extinction probability* $p^{(0)}(x,t)$ is a non-decreasing function of t and converges as $t \to \infty$ to the *asymtotic extinction probability* $p_e(x)$, which is the smallest non-negative solution of the functional equation

$$\xi(x) = G[\xi\,|\,x;t]; \tag{7.33}$$

(2) In the case of multiplicative chains, let $p^{(n)}(x) = P^{(n)}(X^{(n)}\,|\,x;1)$, $m(x) = \sum_{n=1}^{\infty} np^{(n)}(x)$ and $m_{(2)}(x) = \sum_{n=2}^{\infty} n(n-1)p^{(n)}(x)$; then $p_e(x) \equiv 1$ if $\sup_{x\in X} m(x) < 1$ and $\sup_{x\in X} p_e(x) < 1$

if $\inf_{x \in X} m(x) > 1$ and $m_{(2)}$ is bounded; (3) In the continuous time case, if the p.g.fl. of P satisfies (7.22), time-homogeneity implies that $Q(A, t \mid x, s)$ depends only on $t - s$; hence we write it as $Q(A, t \mid x)$. Let $W(A \mid x) = \lim_{t \to \infty} Q(A, t \mid x)$ and let K be the p.g.fl. of W. We define an *associated multiplicative chain* by the sequence of p.g.fls. $\{K_n\}$ defined recursively by $K_{n+1}[\xi \mid x] = K\{K_n[\xi \mid \cdot] \mid x\}$. The asymptotic extinction probabilities of the continuous time process and its associated chain are identical, so that (2) applies with m and $m_{(2)}$ respectively the total mean and 2nd order factorial moment of W.

References

[1]. M. S. BARTLETT, Processus stochastiques ponctuels. *Ann. Inst. Henri Poincaré*, 14 (1954), 35–60.

[2]. ——, *An Introduction to Stochastic Processes*. Cambridge, 1955.

[3]. M. S. BARTLETT & D. G. KENDALL, On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology. *Proc. Cambridge Philos. Soc.*, 47 (1951), 65–76.

[4]. H. J. BHABHA, On the stochastic theory of continuous parametric systems and its application to electron cascades. *Proc. Roy. Soc.* A 202 (1950), 301–322.

[5]. J. L. DOOB, *Stochastic Processes*. New York, 1953.

[6]. W. FELLER, *An Introduction to Probability Theory and its Applications*. 2nd edition, New York 1957.

[7]. T. E. HARRIS, Some mathematical models for branching process. *2nd Berkeley Symposium on Math. Statistics and Probability* (1951), 305–328.

[8]. E. HILLE & R. S. PHILLIPS, *Functional Analysis and Semi-Groups*. Amer. Math. Soc. Colloq. Publications, Vol. 31, 1957.

[9]. D. G. KENDALL, Stochastic processes and population growth. *J. Roy. Statist. Soc. Ser. B*, 11 (1949), 230–264.

[10]. A. N. KOLMOGOROV, *Foundations of the Theory of Probability*. New York, 1950.

[11]. J. E. MOYAL, "Statistical problems in nuclear and cosmic ray physics", *Bull. Inst. International Statistique*, 35 (1957), 199–210.

[12]. ——, Discontinuous Markoff processes. *Acta Math.*, 98 (1957), 221–264.

[13]. J. NEYMAN & E. L. SCOTT, Statistical approach to problems of cosmology,. *J. Roy. Statist. Soc. Ser.* B 20 (1958), 1–43.

[14]. A. RAMAKRISHNAN, Stochastic processes relating to particles distributed in a continuous infinity of states. *Proc. Cambridge Philos. Soc.*, 46 (1950), 595.

[15]. I. E. SEGAL, Abstract probability spaces. *Amer. J. Math.* 76 (1954), 721–732.

[16]. S. ULAM, Zur Mass-Theorie in der allgemeinen Mengenlehre. *Fund. Math.*, 16 (1930), 140–150.

[17]. H. WOLD, Sur les processus stationnaires ponctuels. *Le Calcul des Probabilités et ses Applications*, Publications du C.N.R.S. t. 13, 1949.

Received May 30, 1961, in revised form Nov. 14, 1961