No. 209 (84-4)

A FOUNDATIONAL FRAMEWORK FOR POISSON FREQUENCY ANALYSES OF WEAKLY INTERACTING POPULATIONS

by

Tony E. Smith*

January, 1984

* Professor of Regional Science, University of Pennsylvania, Philadelphia, PA. Currently visiting the Institute of Socio-Economic Planning, University of Tsukuba, Sakura, Ibaraki 305, Japan.

ABSTRACT

An axiomatic foundation is developed for Poisson frequency analyses of population processes. In particular it is shown that for a wide range of population processes, Poisson frequencies can be characterized in terms of two simple independence axioms: (i) independence of the individual states of population members for any given population size, and (ii) independence of the population frequencies within any given partition of states. Hence, for all independent population processes satisfying these two conditions, one must necessarily employ the Poisson distribution for all frequency analyses. More generally, it is shown that for a variety of weakly interacting populations in which these conditions are approximately satisfied, the Poisson distribution continues to provide a natural framework for analysis. In addition, it is shown that this notion of independent population processes yields a new characterization of classical spatial Poisson processes. Finally, these results are applied to a specific class of population processes involving spatial flow phenomena.

1. INTRODUCTION

In the study of large scale population behavior, one rarely has available detailed time series on individual behavior. Rather, researchers are typically confronted with unordered data sets which, at best, reflect the cross sectional character of the given population. Within this restricted framework, however, many properties of population behavior can still be revealed by means of categorical analysés in which population frequencies are studied with respect to partitions of relevant population attributes [as developed in Haberman (1964,1979), Bishop, Fienberg, and Holland (1975), Nishisato (1980), among many others]. For example, in studying the possible causes of a given disease, one may attempt to associate high disease frequencies with certain attributes of individuals. Similarly, in studying urban travel behavior, one may attempt to explain observed trip frequencies in terms of various attributes of both individual trip makers and their destination choices. More generally, one may study variations in mean population frequencies with respect to a number of relevant population attributes.

In this context, it has long been recognized that the assumption of independent Poisson frequencies yields a very convenient statistical framework for analysis [see for example the excellent discussion in Haight (1964) and Plackett (1974)]. In particular, since the Poisson distribution is paramaterized solely in terms of its mean, this distributional assumption allows a wide range of candidate mean-frequency models to be estimated and tested without the introduction of extraneous parameters. Hence for those sampling contexts in which this assumption is appropriate, mean frequencies can be analyzed in a simple and powerful way.

However, when it comes to justifying this Poisson assumption, practitioners tend to be rather vague. For example, it is often loosely asserted that independent frequency totals somehow imply the Poisson distribution, or that Poisson frequencies correspond to random sampling with unconstrained sample size (both of which are false without further assumptions). A somewhat more accurate argument asserts that sample observations occurring independently over time lead to Poisson frequencies. While this assertion still requires further assumptions, it captures the central feature of most Poisson models -- namely their intrinsic dependence on the structure of some underlying continuum over time or space. More specifically, sample observations are modeled as "rare events" within some continuum of possible events [as discussed by Battacharyya and Johnson (1977,section 5.8) and others]. For example, the time-independence argument above involves a temporal stochastic process in which it is postulated that sample frequencies in any time interval depend on the length of that interval, and in particular shrink to zero as the interval becomes small. Similarly, in spatial Poisson processes it is postulated that frequencies in each region of space depend on some underlying measure of volume, and shrink to zero as volumes become small (see section 2.4.1 below). While such assumptions may be very appropriate for modeling certain explicit "counting processes" (such as time counts of telephone calls over a line, or spatial counts of flowers in a field), they hardly qualify as general descriptions of the many types of frequency data to which the Poisson distribution is usually applied. For example, in most large sample surveys (such as National Census data) the actual sampling procedure over space and time is highly complex and difficult to model explicitly. Indeed, even when explicit time sampling is meaningful (as for example in urban traffic counts), the inherent complexity of large scale systems often necessitates the use of partial and relatively unsystematic sample schemes. Hence, the classical assumptions of temporal and spatial Poisson processes are at best artificial in such situations (see section 4.2 for further discussion). Moreover, while alternative formulations of Poisson processes are possible which do not involve a continuum of potential events, the fundamental notion of "rare events" continues to play a central role (see section 4.1 below).

With this in mind, the central purpose of the present paper is to propose an alternative approach to Poisson frequency processes which is entirely free of such rare event concepts, and which instead focuses solely on the statistical independence properties of observable population events. While the possibility of such an approach has essentially been known for twenty years [and follows from the results of Chatterji (1963) as shown in section 3.3 below], it has received surprisingly little attention in practice. Hence our main objective is to formulate this independence approach in a manner which clarifies its applicability to a wide range of population phenomena, and to illustrate its application to "spatial flow" processes in particular (see sections 2.5 and 3.4 below). This formulation is also shown to yield a new characterization of classical spatial Poisson processes (sections 2.4.1 and 3.5 below), and finally, is shown to provide a natural "bench mark" for the analysis of weakly interacting populations in general (section 2.4.2 and section 4 below).

To establish these results within a general setting, we begin by formulating a broad class of population processes consisting essentially of a random finite number of individuals, each occupying states in some abstract state space. With a view toward applications, our only structural assumptions on such processes are that the probable states of the population not depend on the labeling of individual members (so that analysis of unordered data sets is meaningful), and that the underlying state space be partitionable in some nontrivial way (so that categorical analyses are meaningful).

Within this setting, our main result (Theorem 3.1 below) is to show that if the states of individual population members are statistically independent for each given population size, then independent frequency totals do indeed imply the Poisson distribution. In other words, if "state independence" is added to "frequency independence", then these two conditions together characterize Poisson frequencies. From a practical viewpoint, this means that for all independent population processes satisfying these two conditions, one must necessarily employ the Poisson distribution for frequency analyses. More generally, for those weakly interacting populations in which such independence conditions may serve as reasonable hypotheses (as discussed in section 4 below), this result provides a natural basis for employing the Poisson framework.

In order to highlight the essential features of this approach, we begin in the next section with an informal development and discussion of the main results, including an application to spatial flow processes. This is followed in section 3 by a formal development of independent population processes and their consequences. Finally, the application of such processes to more general types of weak interaction behavior is considered in section 4.

2. INFORMAL DEVELOPMENT AND DISCUSSION OF RESULTS

To establish a concrete setting for the analysis of population frequency data, we begin in section 2.1 below with a development of population processes in general. The concept of an independent population process is then motivated in terms of the basic hypotheses of state independence and frequency independence in section 2.2. The major structural consequences of these hypotheses, including the Poisson Frequency Theorem and implied state probability distributions, are presented in section 2.3. Additional theoretical consequences for classical spatial Poisson processes and more geleral types of weak interaction processes are discussed in section 2.4. Finally, these results are applied to the important class of spatial flow processes in section 2.5.

2.1 General Population Processes

Consider a population consisting of a random finite number of individuals, i, each occupying some state, ω_i, in a relevant state space, Ω. Each possible population state may then be represented as a finite sequence $(\omega_1,..,\omega_n)$ consisting of the states of its individual members. For example, in the case of human populations (such as political party members, disease victims, etc.), these states may consist of the relevant attribute profiles (age, sex, occupation,...) for each member. To analyze the observed frequencies of such attributes, one must begin by specifying those subsets (or categories) of states in Ω within which such population frequencies can be measured. This collection of <u>measurable sets</u>, M , will be assumed to contain all individual states $\{\omega\}$ in Ω (and to satisfy the usual conditions for a measurable space over Ω as in section 3.1 below). In this context, one may then define the <u>measurable attributes</u> of a given population to be precisely those measurable partitions of Ω . For example, if A and B $(= \Omega - A)$ denote respectively the sets of all "male" and "female" states in Ω , and if A and B are both in M, then one may identify "sex" to be a measurable attribute of this population.

Next, to develop a probability model of such populations, observe that each measurable set A in M may be equivalently viewed as an <u>individual event</u> in which the state ω_i realized by population member i is an element of A (such as the event "i is male"). More generally, each possible sequence $(A_1,..,A_n)$ of sets in M can be interpreted as the <u>population event</u> in which the realized population has exactly n members, and the individual events A_i occur for each member i , respectively. In this way, the probable realizations of the population itself can also be represented in terms of the sets in M . Finally, each possible probability model of such populations is then representable by a <u>probability measure</u>, P , over these population events, in which $P(A_1,..,A_n)$ denotes the probability that event $(A_1,..,A_n)$ occurs.

Within this probabilistic framework, we are interested in population behavior which can in principle be reflected by the resulting frequency counts of its individual member states with respect to various measurable attributes. This requires that the probable realizations of population states $(\omega_1,..,\omega_n)$ not depend on the particular labels "i" of their individual states ω_i . Hence we shall consider only those populations whose individual members are "interchangeable" in the sense that:

P1. (<u>Interchangeability</u>) <u>The probable occurrence of any population event</u> $(A_1,..,A_n)$ <u>does not depend on the particular labeling of its individual events</u> A_i .

Next, in order that categorical analyses be meaningful, there must be at least two distinct individual events which can actually occur. In other words, there must exist some partition of Ω into measurable sets A_1 and A_2 such that it is possible for individual states to occur in either A_1 or A_2 . Moreover, if population size is subject to no prior restrictions, then one may suppose that any number of states can occur in A_1 or A_2 . Hence we now consider those populations which also satisfy the following "positive divisibility" property:[1]

P2. (Positive Divisibility) There is some partition of Ω into measurable sets A_1 and A_2 for which the probable occurrence of any finite numbers of individual states in A_1 and A_2 is positive.

These two properties together complete our basic probability framework for modeling population frequency data. Hence we now summarize this class of probability models as follows (see Definition 3.1 below for a more precise statement):

DEFINITION 2.1 Each probability measure, P, over population events satisfying properties P1 and P2 is designated as a population process.

To analyze the structure of such processes, it is convenient to define certain random variables by which they can be represented. First of all, observe that if $P^n(A_1,..,A_n)$ denotes the conditional probability of the event $(A_1,..,A_n)$ given that the realized population of size n, then with respect to this conditional probability measure, P^n, it is meaningful to define random state variables, S_i, denoting the states of each individual population member $i = 1,..,n$. The joint distribution of these random variables is then given by the conditional event probabilities themselves, i.e. $P^n(S_1 \in A_1,..,S_n \in A_n) = P^n(A_1,..,A_n)$.

Next, for each measurable set A in Ω, one may define a random frequency variable, N_A, denoting the number of population members with states in A. In particular, the random variable N_A assigns to each population state $(\omega_1,..,\omega_n)$ the number of individual states in $(\omega_1,..,\omega_n)$ which are elements of A. If for any measurable sets $A_1,..,A_k$ we denote by $(N_{A1} = n_1,..,N_{Ak} = n_k)$ the event that each set A_i contains exactly n_i population members, then the joint distribution of N_{Ai}, $i = 1,..,k$, is given by these event probabilities $P(N_{A1} = n_1,..,N_{Ak} = n_k)$. Such joint frequencies constitute the basic data for all categorical analyses. For example, if one is interested in the frequencies of various occupation types $i = 1,..,k$ within some population of workers, and if $A_i \in M$ denotes the measurable set of possible states for workers which include job type i, then one may study the joint distribution of the job frequency variables $N_{A1},..,N_{Ak}$. (Examples involving multiple attribute comparisons are given in sections 2.3.1 and 2.5.2 below).

2.2 Independent Population Processes

As mentioned in the introduction, most analyses of population frequency variables tend to focus mean frequency values. Indeed, all standard structural models of categorical data [including the many models developed in Haberman (1964), Bishop, Fienberg, and Holland (1975), and Nishisato (1980), among others] postulate that the relevant structural interaction within the population of interest can be captured by mean frequency counts. Hence it is implicitly assumed that any additional effects due to interactions or ordering relationships among individuals within that population can be ignored.

The practical consequences of this assumption are best illustrated by means of an example. Consider for instance the problem of identifying those physical attributes of individuals which influence their susceptibility to a given disease, X . Suppose in particular that one is given data on various attributes (age, sex, race,..) of current disease victims, and proceeds to analyze the frequencies of these attributes. As an extreme case, imagine that disease X is highly contagious, and has been introduced into the general population through a particular racial group (say, by recent immigrants to that group). Then it is almost certain that individual contacts among group members will generate a high frequency of within-group victims. Hence, any statistical analysis of such frequencies is bound to suggest that "race" is a critical factor in determining susceptibility to disease X (perhaps leading one to investigate physical differences among racial groups). Here it is clear that strong interaction effects among individuals (in this case contagion effects) tend to overwhelm any real physical susceptibility considerations. While these effects could of course be isolated and accounted for within an explicit "epidemic" modeling framework, such models require detailed time series data for analysis. Hence the point of this illustration is to show that analysis of population frequency data alone is inadequate in many situations, and may even be very misleading.

On the other hand, consider a disease, Y , which is known to be noncontagious (such as heart disease). In this case, it may be quite appropriate to postulate that interactions between individuals play no significant role in determining susceptibility to Y . Hence population frequency data may be very effective in identifying key factors of susceptibility.[2] Moreover, even for many contagious diseases (such as the common cold), exposure levels with the general population are often sufficiently diffuse to allow these interaction effects to be discounted. In other words, if all people are equally exposed to colds, then those more susceptible will still tend to get colds more often. (For further discussion of such diffuse interaction effects see section 4.2 below).

2.2.1 State Independence Hypothesis

More generally, in cases where only weak interactions are present within a given population, it may often be reasonably hypothesized that the relevant attributes of individual members are not significantly influenced by the attributes of other members. This hypothesis can be stated more precisely in terms of the state variables $S_1,..,S_n$ associated with a given population of n members (such as the attribute profiles associated with a current population of disease victims). In particular, if the probable realizations of these state variables are not significantly influenced by one another, then it may be hypothesized that:

H1. (State Independence) For any given population size, n , the state variables $S_1,..,S_n$ of individual members are statistically independent.

2.2.2. Frequency Independence Hypothesis

Next, consider the population members associated with individual states. Again, if population membership is not significantly affected by interactions among individuals, then the number of individuals exhibiting any particular state should not be influenced by the numbers of individuals exhibiting other states.[3] For example, if a disease is noncontagious (or is uniformly contagious throughout the population), then the number of males with the disease should not influence the number of females with the disease.[4] More generally, if no significant interactions among possible states are present in a given population, then it may be hypothesized that:

H2. (Frequency Independence) For any choice of nonoverlapping measurable sets $A_1,..,A_k$ in the state space Ω , the corresponding population frequencies $N_{A1},..,N_{Ak}$ are statistically independent.

This pair of independence hypotheses may thus be said to characterize those populations in which interaction effects among individual members are not significant. By way of summary, we now say that:

DEFINITION 2.2 Each population process, P , consistent with hypotheses H1 and H2 is designated as an independent population process.

2.3. Properties of Independent Population Processes

The single most important property of independent population processes, as we shall demonstrate, is that all population frequencies for such processes are always Poisson distributed. This in turn implies the existence of certain well defined "state probabilities" for such processes, with respect to which observed relative frequencies are the natural maximum likelihood estimators. Each of these results will be discussed in turn.

2.3.1 Poisson Frequency Theorem

For each measurable set A in Ω with associated frequency variable N_A, let $\mu(A) = E(N_A)$ denote the mean population frequency in A. In terms of this notation, the central property of independent population processes can now be stated as follows (see Theorem 3.1 below):

POISSON FREQUENCY THEOREM. If P is an independent population process, then the joint frequency distribution for each measurable partition $\{A_1,..,A_k\}$ of Ω is given for all $n_1,..,n_k$ by

$$(2.1)\qquad P(N_{A_1} = n_1;\ldots, N_{A_k} = n_k) = \prod_{i=1}^{k}\left\{\frac{\mu(A_i)^{n_i}}{n_i!}\exp[-\mu(A_i)]\right\}$$

From a theoretical viewpoint, the most important consequence of this result is to show that Poisson processes can be characterized entirely in terms of simple independence concepts. In particular, it shows that independent population processes must give rise to Poisson frequencies regardless of the nature of their underlying state spaces. Hence, for any population process in which these general independence hypotheses are deemed appropriate, one may analyze unordered frequency data within a Poisson framework without appealing to any specific temporal or spatial sampling structure.

From a practical viewpoint, the most important consequence of the Poisson Frequency Theorem is of course to permit direct estimation and testing of a wide variety of explicit models of mean frequencies. As one simple illustration, consider the relationship between two population attributes, such as between occupations and political orientations of individuals within a given society. If the individual state space Ω is partitioned into n occupation categories $\{O_1,..,O_n\} \subset M$ and into m political party affiliations $\{P_1,..,P_m\} \subset M$, respectively, then by defining the joint partition $\{A_{ij}: i=1,..,n, j=1,..,m\} \subset M$ for all i and j by $A_{ij} = O_i \cap P_j$, one may study this relationship in terms of the two-way contingency table of associated frequency variables $\{N_{ij}: i=1,..,n,\ j=1,..,m\}$. For example, independence of these two attributes can be tested explicitly in terms of the standard multiplicative (log linear) mean frequency model[5]

$$(2.2) \qquad \mu(A_{ij}) \doteq a_i b_j \quad , \qquad i=1,..,n,\ j=1,..,m \quad .$$

By substituting (2.2) into (2.1) and employing standard maximum likelihood techniques, one can estimate and test model (2.2). In this simple case one obtains maximum likelihood estimates $a_i^* = N_{i.} / a^*$ and $b_j^* = N_{.j} / b^*$ where $N_{i.} = \Sigma_j N_{ij}$, $N_{.j} = \Sigma_i N_{ij}$, and $a^* b^* = N = \Sigma_{ij} N_{ij}$. Hence the products $a_i^* b_j^* = N_{i.} N_{.j} / N$ yield a unique maximum value of the likelihood function which may be used to test model (2.2) by standard likelihood ratio procedures. One may also analyze structural interactions explicitly within this framework by employing various parametric mean frequency hypotheses. In the present illustration, one may hypothesize for example that the number n_{ij} of bills supported by party j favoring occupation group i has a positive interaction effect on the composition of party membership. Hence one may postulate a parametric mean frequency model of the form

$$(2.3) \qquad (A_{ij}) = a_i b_j \theta^{n_{ij}} \quad , \qquad i=1,..,n,\ j=1,..,m$$

and test for θ-values differing from unity by substituting (2.3) into (2.1) and employing the same procedures as above. Additional examples of parametric mean frequency models are given in section 2.5.2 below. [For additional details of such maximum likelihood procedures, see for example Haberman (1974, Chapter 2) and Bishop, Fienberg, and Holland (1975, Chapter 3)].

2.3.2 State Probability Measures

A second important property of independent population processes relates to the distribution of the state variables $(S_1,..,S_n)$ for any given population size n . In particular, if N_Ω denotes the total size of the realized population (i.e. the frequency variable for Ω itself), and if $\mu(\Omega) = E(N_\Omega) > 0$ denotes the associated mean population size, then one may readily verify (as in Corollary 3.1 below) that for each population size n and each state variable S_i , $i=1,..,n$,

$$(2.4) \qquad P^n(S_i \in A) = \frac{\mu(A)}{\mu(\Omega)}$$

holds identically for each measurable set A in Ω with mean population frequency $\mu(A)$. Hence, the probability that any individual i from a population of size n occupies a state in A is seen to be the same for all members of finite population realizations, and in particular is always proportional to the mean frequency $\mu(A)$ in A . This suggests the following definition:

DEFINITION 2.3 For each independent population process, P , the unique probability measure, p , defined for each $A \in M$ by

$$(2.5) \qquad p(A) = \frac{\mu(A)}{\mu(\Omega)}$$

is designated as the state probability measure for P .

These state probabilities are also expressible directly in terms of the population process measure, P , as follows (see Corollary 3.1 below):

$$(2.6) \qquad p(A) = P^1(S_1 \in A) = \frac{P(A)}{P(\Omega)}$$

In other words, p(A) is equivalent to the probability that A will be occupied by the single member in a population realization of size $n = 1$.

From a sampling viewpoint, these state probabilities are seen to imply that the conditional state variables $S_1, \ldots, S_n$ for populations of size n are not only independently distributed (as asserted in H1), but are also identically distributed with common distribution (2.5) [or equivalently (2.6)] independent of n . In these terms, the Poisson Frequency Theorem is also seen to suggest a rigorous formulation of the intuitive notion (mentioned in the introduction) that Poisson frequencies correspond to "independent random sampling without fixed sample size". For if hypotheses H1 and H2 are adopted as a formal model such a sampling procedure, then the Poisson Frequency Theorem shows that these samples must always give rise to Poisson frequencies.[6]

As a final corollary of this independent random sampling interpretation, it is well known that Poisson frequencies reduce to the classical multinomial sampling scheme for each fixed sample size. In particular, it follows easily from (2.1) and (2.5) [see Corollary 3.1 below] that for each measurable partition $\{A_1, \ldots, A_k\}$ of Ω , the conditional distribution of the associated population frequencies $N_{A1}, \ldots, N_{Ak}$ for any given fixed population size n is multinomially distributed in terms of the state probabilities $p(A_1),..,p(A_k)$, i.e. that for all $n_1,..,n_k$ with $\sum_i n_i = n$,

$$(2.7) \qquad P^n(N_{A_1} = n_1, .., N_{A_k} = n_k) = n! \prod_{i=1}^{k} \left[\frac{p(A_i)^{n_i}}{n_i!} \right]$$

Hence, within this conditional framework, the relative population frequencies $(n_i/n : i = 1,..,n)$ are seen to be the natural maximum likelihood estimators of these state probabilities, and thus constitute sufficient statistics for estimating and testing any structural models of these probabilities. For example, the multiplicative model of state probabilities corresponding to (2.3) above is seen to have the "logit" form,

$$(2.8) \qquad p(A_{ij}) = \frac{a_i b_j \theta^{n_{ij}}}{\sum_{kh} a_k b_h \theta^{n_{kh}}}$$

and can be estimated and tested by substituting (2.8) into (2.7) and employing maximum likelihood procedures within this multinomial framework. [An interesting comparative analysis of the unconditional Poisson and conditional multinomial approaches is given in Bishop, Fienberg, and Holland (1975, Section 13.4.4)].

2.4 Theoretical Implications for Population Processes

Aside from providing a simple independence characterization of Poisson frequencies, the notion of population processes has a number of broader theoretical consequences. First, it provides a new view of classical spatial Poisson processes. Second, it provides a natural point of departure for the more general study of weakly interacting population processes. Each of these consequences will be discussed in turn.

2.4.1 Homogeneous Spatial Poisson Processes

As a first additional implication of the Poisson Frequency Theorem, one obtains an interesting generalization of the usual notion of a homogeneous spatial Poisson process [as developed, for example, in Cox and Isham (1980, Section 6.2) and Karlin and Taylor (1981, Section 16.1)]. The basic idea of these processes is to construct a natural model of the notion of a "purely random" scattering of points in some space. The classical model focuses on euclidean n-dimensional space, R^n , and considers the point frequencies N_A which might be observed under these conditions in each measurable region A in R^n with volume v(A). If point locations are purely random, then these point frequencies should satisfy the following conditions. First, since the scattering of points should behave the same way everywhere in R^n, point frequencies should satisfy the homogeneity condition that regions of equal volume have the same frequency distributions, i.e. that N_A and N_B should be identically distributed in all regions A and B with v(A) = v(B) . Next, since the numbers of points in disjoint regions should not influence one another, point frequencies should also satisfy the frequency independence condition (H2 above). Finally, in view of the continuum of possible point locations, point frequencies should satisfy the rare-event conditions that (i) regions of very small volume should have almost no chance of containing any points, and that (ii) if they do contain points, they should almost never contain more than one point [see conditions (3.23) and (3.24) below for a precise statement]. Given this formal model of pure randomness, it can be shown [see for example Karlin and Taylor (1981, Theorem 16.1.1)] that the resulting point frequencies N_A in each measurable region A with positive finite volume v(A) are Poisson distributed with mean frequencies $\mu(A)$ proportional to v(A), i.e. satisfying $\mu(A) = \lambda v(A)$ for some constant intensity parameter $\lambda > 0$. While this basic formulation can be extended to spaces in which no volume measure is defined (see the discussion in section 3.3.1 below), all such approaches rely heavily on the basic notion of "rare events".

In this light, the present results suggest a new view of random point patterns. In particular, if together with frequency independence (H2), we interpret "pure randomness" to mean that the locations of points in any finite pattern of n points should not influence one another, i.e. state independence (H1), then subject only to the mild structural conditions P1 and P2 (i.e. that point patterns not depend on the labeling of individual points, and that the space Ω of possible locations have some nontrivial partition into distinct regions) one can conclude from the Poisson Frequency Theorem that point frequencies must continue to be Poisson distributed. Moreover, since the mean frequencies $\mu(A)$ for each measurable set A in M are easily seen to define a finite additive measure, μ, over the state space Ω, one may reinterpret this theorem to assert that for each "purely random" point process satisfying H1 and H2 there exists a "volume measure" on Ω (namely the mean frequency measure, μ) with respect to which point frequencies are representable by a homogeneous spatial Poisson process (with intensity $\lambda = 1$). Within this broader framework, the classical notion of purely random point patterns can be extended to "lumpy" spaces (such as finite spaces) in which the notion of "small sets" is not appropriate, and more generally, may be extended to cases in which no notion of "rare events" is meaningful.

Finally, when there does exist a meaningful notion of volume in a given state space, the above results yield a new characterization of homogeneous spatial Poisson processes for this case. In particular, if one adds to H1 and H2 the weaker homogeneity condition that only mean frequencies depend on volume, i.e. that

H3. (Homogeneity) All measurable sets with the same volume have the same mean point frequencies.

then population processes satisfying (H1,H2,H3) always generate homogeneous spatial Poisson processes (see Theorem 3.2 below).

2.4.2 A Bench Mark Model for Weak Interaction Behavior

To illuminate the broader implications of the Poisson Frequency Theorem, it is instructive to interpret this result within the general classification of large scale systems proposed by Weaver (1958), and elaborated by Wilson (1974,1977) in the context of social systems. In particular, these authors distinguish between two possible types of complexity in large scale systems, designated respectively as organized complexity and disorganized complexity. The essential difference between the two can be illustrated in the context of social systems by contrasting governmental decisions with various types of individual decision behavior within a given society. For example, a single budget allocation decision by government can profoundly influence the subsequent actions of large numbers of individuals or groups of individuals within that society. On the other hand, the budget allocation decision of any individual member, say a household, generally has only a small influence on the behavior of other members. Hence, the former type of decision behavior may be said to involve strong interaction effects, while the latter type involves only weak interaction effects among behaving units.

Such distinctions between weak and strong interactions arise in many other contexts as well. For example, in the illustration of disease populations given in section 2.2 above, a distinction was made between contagious and noncontagious diseases. Here the introduction of a new contagious disease into a population (or even the first cold of the season) was seen to result in strong interaction effects among the states of individuals. More generally, such distinctions between weak and strong interaction effects can often be associated with equilibrium versus nonequilibrium states of the system. In particular, the global changes resulting from the introduction of a new disease into a given population (or a new product into a given economy) can be described as temporary nonequilibrium states of the system, leading usually to some new equilibrium state. On the other hand, population behavior involving only weak interactions (such as diffuse exposure to colds during the height of the cold season) may be said to involve equilibrium behavior in the sense that such interactions lead to no global changes in the system.[7] (See sections 4.2 and 4.3 below for further discussion of these ideas.)

While it is difficult to define the above concepts in a completely satisfactory way, it may nonetheless be seen intuitively that population behavior involving only weak interactions is more amenable to statistical averaging than behavior involving strong interaction effects. For example, in analyzing various governmental actions, it is clear that small numbers of decision makers, and even individual personalities themselves, can significantly influence aggregate system behavior. However, in analyzing patterns of household purchases, for example, one can often ignore detailed distinctions between households, and still obtain a good picture of overall system behavior by averaging individual household purchase decisions. More generally, it may be argued [as in Weaver (1958) and Wilson (1974,1977)] that the behavior of large populations involving only weak interactions among individual members can be well described by statistical averaging.

In this light, the Poisson Frequency Theorem suggests at least one way in which these intuitive notions can be made precise. For if weak interaction behavior within populations is formalized in terms of the statistical independence hypotheses H1 and H2 , then this theorem is seen to give rigorous meaning to the idea that such population behavior can be well described by system averages. Indeed, since the Poisson distribution is solely parameterized by its mean, it follows at once that all stochastic behavior in such populations is completely characterized by mean population frequencies. More generally, it may be argued that to the extent that a given population process is consistent with H1 and H2, behavior within that population should be well represented by average behavior. In this sense, the formal model of an independent population process may be said to provide a natural bench mark for the study of weakly interacting populations in general. (These ideas will be considered in more detail in the final section of this paper.)

To illustrate the application of independent population processes to more concrete situations, we now apply this framework to population behavior involving "spatial flows". For a broader range of possible applications, see for example the many illustrations in Haberman (1964), Bishop, Fienberg, and Holland (1975) and Nishisato (1980).

2.5 Application to Spatial Flow Processes

Spatial flows are involved in a wide variety of population behavior. For example, migration behavior, commuting behavior, and shopping behavior all involve flows of people over space. Similarly, communications involve flows of information over space, and economic transactions involve flows of goods and services over space. Moreover, in large systems of such spatial flows, one often observes only weak interactions between individual flows. For example, daily decisions on where to shop or whom to visit are usually specific to individuals or households. Even major decisions such as where to live, or when and where to move, are often only weakly linked to the decisions of other individuals or households. More generally, as long as the system of flows being observed is assumed to be near equilibrium (i.e. not undergoing major structural changes), one may hypothesize that only weak interactions among flows are present.[8] (See also section 4.3 below.)

2.5.1 Independent Flow Processes

In this context, it is appropriate to model such flow systems statistically in terms of independent population processes, where the relevant "population" now consists of the individual spatial flows (trips, purchases, telephone calls) occurring during some given period. To be more specific, we begin by postulating a space, X, of potential flow origins, x, and a space, Y, of potential flow destinations, y. For example, in the case of urban shopping trips, X may be thought of as a map giving the locations, x, of potential shoppers in an urban area, and similarly, Y may be a map giving the locations, y, of all potential shopping opportunities in the area. A flow is then described by an origin-destination pair (x,y) and the relevant state space Ω consists of all possible origin-destination pairs from X to Y.

Within this context, one can then model flow behavior as a population process with population states corresponding to flow patterns $((x_1,y_1),..,(x_n,y_n))$ in Ω. Postulates P1 and P2 for the resulting flow process, P, correspond respectively to the assumptions that (i) the realization of any given flow pattern does not depend on the particular ordering of its individual flows, and that (ii) the underlying space of origins and/or destinations is partitionable into at least two nontrivial zones or subregions. (See section 3.4 below for a more precise statement.)

The additional assumptions required for an independent flow process, P, amount to a formalization of the hypothesis that no strong interactions among flows are present. In particular, H1 asserts that for any flow pattern $((x_1,y_1),..,(x_n,y_n))$ of size n, the individual flows (x_i,y_i) are statistically independent of one another, and

H2 asserts that the flow frequencies in any nonoverlapping sets of possible origin-destination pairs in Ω be statistically independent. (See section 3.4 below for a more precise statement.) For those flow contexts in which these hypotheses are appropriate, the Poisson Frequency Theorem provides a powerful analytical framework within which a wide variety of specific flow models can be estimated and tested. The many advantages of this Poisson framework have recently been recognized by a number of researchers in spatial flow modeling, including Sen and Sööt (1981), Flowerdew and Aitkin (1982), and Gray and Sen (1983). Hence the present concept of independent flow processes can usefully be regarded as a possible formal foundation for such studies.

2.5.2 Gravity Models of Flow Frequencies

To illustrate the application of this framework to specific flow models, it is convenient to consider a measurable partition of the origin space X into a finite number of origin zones $\{X_i : i \in I\}$, and similarly, a partition of Y into destination zones $\{Y_j : j \in J\}$. As in section 2.3.1 above, we may then consider the joint partition $\{\Omega_{ij} : i \in I,\ j \in J\}$ of Ω in which each Ω_{ij} consists of all flows from X_i to Y_j. In this context, if N_{ij} denotes the flow-frequency variable for Ω_{ij} (i.e. the number of flows from X_i to Y_j), and if $T_{ij} = E(N_{ij})$ denotes the associated mean flow frequency, then the Poisson Frequency Theorem implies that for each independent flow process, P, the joint distribution of these flow frequencies (N_{ij}) can be written in terms of mean flow frequencies (T_{ij}) for all possible frequency realizations (n_{ij}) as follows:

$$(2.9) \qquad P(N_{ij} = n_{ij} : i \in I,\ j \in J) = \prod_{i \in I} \prod_{j \in J} \left[\frac{T_{ij}^{\,n_{ij}}}{n_{ij}!} \exp(-T_{ij}) \right]$$

Hence, for any given parametric model of mean flow frequencies, (2.9) provides an explicit finite-parameter family of distributions within which standard maximum likelihood estimation and testing procedures can be employed.

As a specific example, one may consider the generalized gravity model of mean flow frequencies studied by Sen and Sööt (1981) and Gray and Sen (1983). Here it is hypothesized that mean flow frequencies are exponential functions of some set of separation measures $\{c_{ij}^k : k \in K\}$ between origin zones, i, and destination zones, j, (where, for example, c_{ij}^1 = travel cost between i and j, c_{ij}^2 = travel time between i and j, and so on).[9] In particular, it is hypothesized that there exist positive origin weights $a = (a_i : i \in I)$, destination weights $b = (b_j : j \in J)$, and separation *weights* $\theta = (\theta_k : k \in K)$ such that the mean flow frequencies are expressible for each $i \in I$ and $j \in J$ as

(2.10) $$T_{ij} = T_{ij}(a,b,\theta) = a_i b_j \exp(\textstyle\sum_k \theta_k c_{ij}^k)$$

Substitution of (2.10) into (2.9) then yields an explicit distribution for estimating and testing this finite-parameter family of exponential models. [For a theoretical development of the parametric maximum likelihood methods appropriate for estimating and testing such models, see Haberman (1964, Chapter 2). With respect to model (2.10) in particular, computational procedures are developed and applied in Sen and Sööt (1981) and Gray and Sen (1983)].

Alternatively, if one is only interested in modeling *relative* flow frequencies among zones, then the multinomial distribution in (2.7) provides the appropriate framework for analysis. In particular, it follows [as in (2.8) above] that if we now define the *flow probabilities*, p_{ij}, corresponding to the mean frequencies, T_{ij}, in (2.10) by

(2.11) $$p_{ij} = p_{ij}(a,b,\theta) = \frac{a_i b_j \exp(\sum_k \theta_k c_{ij}^k)}{\sum_{s \in I} \sum_{t \in J} a_s b_t \exp(\sum_k \theta_k c_{st}^k)}$$

then for any given total number of flows, n, the specific flows observed can be treated as an independent random sample of size n from this probability distribution. Hence, by substituting (2.11) into (2.7), we obtain an explicit multinomial sampling distribution of the form:

(2.12) $$P^n(N_{ij} = n_{ij} : i \in I, j \in J) = n! \prod_{i \in I} \prod_{j \in J} \left[\frac{p_{ij}(a,b,\theta)^{n_{ij}}}{n_{ij}!} \right]$$

A specific computational comparison of this conditional multinomial approach [(2.11), (2.12)] with the unconditional Poisson approach [(2.9),(2.10)] is given in Sen and Sööt (1981). Additional examples of such parametric models are given in Haberman (1979), among others.

3. FORMAL DEVELOPMENT AND ANALYSIS

To establish the above results in a precise way, we begin with a formal definition of population processes in section 3.1 below. This development follows the basic framework of Moyal (1962). [A more abstract treatment of essentially the same framework is given in Carter and Prenter (1972)]. Throughout the analysis we shall employ the following standard notatation. Let Z_o and Z_+ denote the nonnegative integers and positive integers, respectively. Let $A_1 \times \cdots \times A_k$ denote the product of nonempty sets $A_1, .., A_k$, and for each $k \in Z_+$ let A^k denote the k-fold product of set A with itself. For notational simplicity, we adopt the convention that $A^o \times B = B \times A^o = B$ for all sets A and B.

3.1 Population Processes

Consider an arbitrary set Ω of (at least two) individual states, ω, and let M denote a fixed σ-algebra of individual events in Ω containing all singleton sets $\{\omega\} \subset \Omega$. Given this measurable space (Ω, M) of individual states, each n-tuple $\omega^n = (\omega_1, .., \omega_n) \in \Omega^n$, $n \in Z_+$, then represents a possible population state for a population of n individuals in (Ω, M). Hence, the relevant class of population events is given by the σ-algebra M^n of sets in Ω^n generated by M (i.e. the unique smallest σ-algebra of sets containing all products $A_1 \times \cdots \times A_n$ of individual events in M). If 0 denotes the null population state for a population of zero members, then by setting $\Omega^o = \{0\}$ and $M^o = \{\emptyset, \Omega^o\}$, respectively, we may define the universe of all finite population states to be $\Omega^* = \cup_{n \in Z_o} \Omega^n$, and take the associated class of all finite population events to be the σ-algebra M* generated by the population events in $\cup_{n \in Z_o} M^n$. Hence, the relevant measurable space of all finite population states is given by (Ω^*, M^*). Finally, each probability measure, P, on (Ω^*, M^*) with $P(\Omega^*) = 1$ defines a probability space (Ω^*, M^*, P) over the possible finite population realizations in Ω^*.

To analyze the associated frequency distribution of realized individual states in Ω, let the indicator function $\delta_A: \Omega \to \{0,1\}$ associated with each measurable set $A \in M$ be defined for each $\omega \in \Omega$ by $\delta_A(\omega) = 1$ for $\omega \in A$ and $\delta_A(\omega) = 0$ otherwise. Then for each population state $\omega^n = (\omega_1, .., \omega_n) \in \Omega^*$, the number of individual states ω_i in A is given by:

$$(3.1) \qquad N_A(\omega^n) = \sum_{i=1}^{n} \delta_A(\omega_i)$$

The function $N_A: \Omega^* \to Z_o$ generated by (3.1) is M*-measurable [Moyal (1962, Section 3)] and hence defines a random frequency variable, N_A , on (Ω^*, M^*, P) associated with the set A . The resulting family of random variables $N = (N_A: A \in M)$ is designated as the frequency process on (Ω, M) generated by (Ω^*, M^*, P).[10] Such processes are in turn describable by their finite-dimensional distributions associated with each finite measurable partition of Ω . More precisely, if each family of sets $\{A_1,..,A_n\} \subseteq M$ satisfying $\Omega = \cup_{i=1}^{k} A_i$ and $A_i \cap A_j = \emptyset$ for $i \neq j$ is designated as a measurable k-partition of Ω , then every frequency process N on (Ω, M) is uniquely defined by the joint distributions of $(N_{A1},..,N_{Ak})$ for all measurable k-partitions of Ω with $2 \leq k < \infty$ [Moyal (1962, Theorem 3.2)]. To specify these joint distributions, observe that for any $n_1,..,n_k \in Z_o$ with $n = \sum_{i=1}^{k} n_i$, the event $(N_{A1} = n_1,.., N_{Ak} = n_k)$ occurs iff each of the individual events A_i occurs exactly n_i times, i.e. iff some permutation (relabeling) of the individual event sequence $A_1^{n_1} \times \cdots \times A_k^{n_k} \in M^n$ occurs. Hence, if Π_n denotes the set of all permutations $\pi = (\pi_1,..,\pi_n)$ of the integers (1,..,n), and if for each population state $\omega^n = (\omega_1,..,\omega_n) \in \Omega^n$ we let $\omega_\pi^n = (\omega_{\pi_1},..,\omega_{\pi_n}) \in \Omega^n$ denote the relabeling of ω^n defined by π , then the associated relabeling of each event $A \in M^n$ is given by $A_\pi = \{\omega_\pi^n : \omega^n \in A\}$. In these terms, the event $(N_{A1} = n_1,..,N_{Ak} = n_k)$ is seen to correspond to the event $\cup_{\pi \in \Pi_n} (A_1^{n_1} \times \cdots \times A_k^{n_k})_\pi \in M^n$, so that the joint distribution of $(N_{A1},..,N_{Ak})$ is given by:

$$(3.2) \qquad P(N_{A_1} = n_1,..., N_{A_k} = n_k) = P[\cup_{\pi \in \Pi_n} (A_1^{n_1} \times \cdots \times A_k^{n_k})_\pi]$$

Given these definitions, we now restrict our analysis to stochastic population behavior which is completely describable by its associated frequency process N . This is equivalent [Moyal (1962, Theorem 3.1)] to the requirement that the probable occurrence of any event $A \in M^n \subseteq M^*$ be invariant with respect to the labeling of the n population members, i.e. that $P(A) = P(A_\pi)$ hold for all $\pi \in \Pi_n$. Moreover, to avoid degenerate cases, it is also assumed that there exist nontrivial partitions of Ω . In particular, it is enough to assume that there exists some measurable 2-partition $\{A_1, A_2\}$ of Ω for which the associated joint frequency distribution (N_{A1}, N_{A2}) is everywhere positive. Hence, as a formalization of Definition 2.1 above, we now have:[11]

DEFINITION 3.1 A probability measure P on (Ω^*, M^*) is designated as a population process on Ω^* iff P satisfies the following two conditions for all $n_1, n_2 \in Z_o$, $n \in Z_+$, $\pi \in \Pi_n$, $A \in M^n$, and for some 2-partition $\{A_1, A_2\} \subseteq M$,

P1. (Interchangeability) $\qquad P(A) = P(A_\pi)$

P2. (Positive Divisibility) $\qquad P(N_{A_1} = n_1, N_{A_2} = n_2) > 0$

Observe in particular that condition P2 ensures that each finite population size is possible, i.e. that $P(N_\Omega = n) = P(\Omega^n) > 0$ for all $n \in Z_o$.[12] Hence the <u>conditional probability measure</u> P^n on (Ω^n, M^n) defined for all $A \in M^n$ by $P^n(A) = P(A) / P(\Omega^n)$ yields a well defined conditional probability space (Ω^n, M^n, P^n) for each $n \in Z_+$. Within this conditional framework, the family of M^n-measurable functions $S_i : \Omega^n \rightarrow \Omega$, $i = 1,..,n$, defined for all $\omega^n = (\omega_1,..,\omega_n) \in \Omega^n$ by $S_i(\omega^n) = \omega_i$, yield random <u>state variables</u> S_i on (Ω^n, M^n, P^n) denoting the realized states of each member, i , of this population. The joint distribution of these random variables is given for all $A_1,..,A_n \in M$ by

(3.3) $$P^n(S_1 \in A_1, \ldots, S_n \in A_n) = P^n(A_1 \times \cdots \times A_n)$$

and the associated marginal distribution of each S_i is given for all $A \in M$ by

(3.4) $$P^n(S_i \in A) = P^n(\Omega^{i-1} \times A \times \Omega^{n-i})$$

3.2 Independent Population Processes

Within this general framework, the notion of an independent population process stated in Definition 2.2 above can now be formalized as follows:

<u>DEFINITION 3.2</u> A population process P on (Ω^*, M^*) is designated as an <u>independent population process</u> iff P satisfies the following two additional conditions for all $n, k \in Z_+$, $n_1,..,n_k \in Z_o$, $B_1,..,B_n \in M$, and measurable k-partitions $\{A_1,..,A_k\} \subseteq M$,

H1. (<u>State Independence</u>) $$P^n(S_1 \in B_1, \ldots, S_n \in B_n) = \prod_{i=1}^{n} P^n(S_i \in B_i)$$

H2. (<u>Frequency Independence</u>) $$P(N_{A_1} = n_1, \ldots, N_{A_k} = n_k) = \prod_{i=1}^{k} P(N_{A_i} = n_i)$$

Hence, P is an independent population process on (Ω^*, M^*) if and only if (i) the state variables $S_1,..,S_n$ are independently distributed for each $n \in Z_+$, and (ii) the the frequency variables $N_{A1},..,N_{Ak}$ are independently distributed for each measurable k-partition $\{A_1,..,A_k\}$ of Ω .

3.3 Poisson Frequency Theorem

For each population process P on (Ω^*, M^*), the associated mean frequency measure, μ, is defined on (Ω, M) for each event $A \in M$ with population frequency variable N_A by [see also Moyal (1962, Lemma 3.3)]:[13]

$$(3.5) \qquad \mu(A) = E(N_A) = \sum_{n \in Z_o} n P(N_A = n)$$

Note that while $0 \leq \mu(A) \leq \infty$ is generally possible for each $A \in M$, the positive divisibility condition (P2) implies in particular that $\mu(\Omega) > 0$ must hold for each population process P. In terms of this mean frequency measure, one may now define the notion of a Poisson frequency process as follows:

DEFINITION 3.3 A population process P on (Ω^*, M^*) with finite population mean $\mu(\Omega) < \infty$ is said to generate a (unique) Poisson frequency process, N, on (Ω, M) iff for each measurable k-partition $\{A_1, \ldots, A_k\}$ of Ω and all $n_1, \ldots, n_k \in Z_o$,

$$(3.6) \qquad P(N_{A_1} = n_1, \ldots, N_{A_k} = n_k) = \prod_{i=1}^{k} \left\{ \frac{\mu(A_i)^{n_i}}{n_i!} \exp[-\mu(A_i)] \right\}$$

In this context, our main result is to show that each independent population process P exhibits this Poisson property. The key result from which this property follows is a characterization of the Poisson distribution due to Chatterji (1963). [Related results can be found in Patil and Seshadri (1964), Volodin (1965), and Kagan, Linnik, and Rao (1973, Theorem 13.4.4).] Because of its central importance for the present analysis, it is instructive to give a proof of this result (which slightly simplifies Chatterji's original argument). If a random variable N is designated as integer-valued whenever $P(N \in Z_o) = 1$, then:

LEMMA 3.1 (Chatterji) For any independent integer-valued random variables N_1 and N_2, if for each $n \in Z_+$ the conditional probability of N_1 given $N_1 + N_2 = n$ is positive binomially distributed, i.e. if for all $k = 0, 1, .., n$,

$$(3.7) \qquad P(N_1 = k \mid N_1 + N_2 = n) = \frac{n!}{k!\,(n-k)!} p_n^k (1 - p_n)^{n-k} > 0$$

then $p_n = p_1$ for all $n \in Z_+$, and both N_1 and N_2 are Poisson distributed with positive finite means μ_1 and $\mu_2 = \mu_1(1-p_1)/p_1$, respectively, i.e.

(3.8) $$P(N_i = n) = \frac{\mu_i^n}{n!}\exp(-\mu_i) \quad , \qquad n \in Z_o \ , \ i = 1,2 \ .$$

Proof: The positivity of (3.7) and independence of N_1 and N_2 ensure that the marginal probabilities $P(N_1 = n) = f(n)$ and $P(N_2 = n) = g(n)$ are positive and satisfy

(3.9) $$f(k)\,g(n-k) = P(N_1 = k \mid N_1 + N_2 = n)\,P(N_1 + N_2 = n)$$

for all $k \leq n \in Z_+$. Hence, substituting (3.7) into (3.9), and taking the ratio of $f(k)g(n-k)$ and $f(k-1)g(n-k+1)$, one obtains

(3.10) $$\frac{f(k)\,g(n-k)}{f(k-1)\,g(n-k+1)} = a_n\,\frac{(n-k+1)}{k} \quad , \qquad k = 1,..,n$$

where $a_n = p_n/(1-p_n)$, $n \in Z_+$. Next, by setting $k = n$ in (3.10) and letting $\theta = g(1)/g(0)$, we are led to the following recursion formula for f :

(3.11) $$f(n) = \frac{a_n\,\theta}{n}\,f(n-1) \quad , \qquad n \in Z_+$$

Similarly, by setting $k = 1$ in (3.10) and noting from (3.11) that $f(1)/f(0) = a_1\theta$, we have the following recursion formula for g :

(3.12) $$g(n) = \frac{a_1\,\theta}{n\,a_n}\,g(n-1) \quad , \qquad n \in Z_+$$

To simplify these expressions further, observe that by evaluating (3.10) at $2 = k \leq n$ and employing (3.11) and (3.12), we obtain

(3.13) $$a_n\,\frac{(n-1)}{2} = \frac{f(2)}{f(1)} \cdot \frac{g(n-2)}{g(n-1)} = \frac{a_2\,\theta}{2} \cdot \frac{(n-1)\,a_{n-1}}{a_1\,\theta}$$

$$\Rightarrow \quad a_n = (a_2/a_1)\,a_{n-1} \quad , \qquad n \geq 2$$

which yields the relation $a_n = (a_2/a_1)^{n-1}a_1$ for all $n \in Z_+$. But if $a_2 > a_1$, then by (3.11) we would have $f(n)/f(n-1) = (a_1\theta/n)(a_2/a_1)^{n-1} \to \infty$, which would contradict $1 = P(N_1 \in Z_o) = \sum_n f(n)$. Similarly, if $a_1 > a_2$, then by (3.12), $g(n)/g(n-1) = (\theta/n)(a_1/a_2)^{n-1} \to \infty$ would contradict $1 = P(N_2 \in Z_o) = \sum_n g(n)$. Thus we must have $a_1 = a_2$, and may conclude that $a_n = a_1$ (and hence that $p_n = p_1$) holds identically for all $n \in Z_+$. Finally, by substituting this identity into the recursion relations (3.11 and (3.12), we obtain the following solutions for f and g , respectively:

$$(3.14) \qquad f(n) = \frac{(a_1\theta)^n}{n!} f(0) \quad , \qquad n \in Z_o$$

$$(3.15) \qquad g(n) = \frac{\theta^n}{n!} g(0) \quad , \qquad n \in Z_o$$

which, together with $\sum_n f(n) = \sum_n g(n) = 1$, also imply that $f(0) = \exp(-a_1\theta)$ and $g(0) = \exp(-\theta)$. Hence (3.8) must hold with $\mu_2 = \theta$ and $\mu_1 = p_1/(1-p_1)\theta$, and the result is established. End of Proof.

In addition to this result, we require certain elementary properties of the Poisson distribution. First, the sum of independent Poisson variates $N_1,..,N_k$ with means $\mu_1,..,\mu_k$ is well known to be Poisson distributed with mean $\mu_1 + \cdots + \mu_k$. The converse of this result was first established by Raikov (1938), and will be stated without proof [see for example Lukacs (1970, Theorem 8.8.2)]:

LEMMA 3.2 (Raikov) *If $N_1,..,N_k$ are independent integer-valued random variables with Poisson distributed sum $N_1 + \cdots + N_k$, then each N_i is also Poisson distributed.*

With these preliminary results, we are now ready to establish the main property of independent population processes:

THEOREM 3.1 (Poisson Frequency Theorem) *Each independent population process P on (Ω^*, M^*) generates a unique Poisson frequency process N on (Ω, M) .*

Proof: To establish this result, we first observe that it is enough to show that there exists some measurable 2-partition $\{A_1, A_2\}$ of Ω with Poisson frequencies N_{A1} and N_{A2} (and hence with finite means). For if so, then since N_{A1} and N_{A2} are independent by H2, it will follow that $N_\Omega = N_{A1} + N_{A2}$ must also be Poisson distributed with finite mean $\mu(\Omega) = \mu(A_1) + \mu(A_2) < \infty$. Finally, since H2 also implies that each measurable k-partition $\{A_1,..,A_k\}$ must yield independent frequencies $N_{A1},..,N_{Ak}$ satisfying $N_\Omega = N_{A1} + \cdots + N_{Ak}$, it will then follow at once from Lemma 3.2 that (3.6) must hold identically for all measurable k-partitions of Ω . To establish the existence of such a 2-partition, recall from the positive divisibility property (P2) of population processes that there is some 2-partition $\{A_1,A_2\}$ with $P(N_{A1} = n_1, N_{A2} = n_2) > 0$ for all $n_1, n_2 \in Z_o$. Hence the conditional distribution of N_1 given $N_1 + N_2 = n$ is well defined and positive for all $n \in Z_+$, and it suffices from Lemma 3.1 together with H2 to show that this conditional distribution satisfies (3.7) for some $p_n \in (0,1)$. To do so, observe from (3.4) together with the interchangeability property (P1) of population processes that for all $A \in M$, $n \in Z_+$, and $i = 1,..,n$,

$$
\begin{aligned}
(3.16) \qquad P^n(S_i \in A) &= P(\Omega^{i-1} \times A \times \Omega^{n-i}) / P(\Omega^n) \\
&= P(A \times \Omega^{n-1}) / P(\Omega^n) \\
&= P^n(S_1 \in A)
\end{aligned}
$$

Similarly, since P1 also implies that $P(A_1{}^k \times A_2{}^{n-k}) = P[(A_1{}^k \times A_2{}^{n-k})_\pi]$ holds for all $k \le n \in Z_+$ and $\pi \in \Pi_n$, and since there are only $n!/k!(n-k)!$ distinct relabelings of the set $A_1{}^k \times A_2{}^{n-k}$, it follows from (3.2) that [see also (3.5) in Moyal (1962)]:

$$
\begin{aligned}
(3.17) \qquad P(N_1 = k \mid N_1 + N_2 = n) &= P^n(N_1 = k, N_2 = n-k) \\
&= P[\cup_{\pi \in \Pi_n} (A_1{}^k \times A_2{}^{n-k})_\pi] / P(\Omega^n) \\
&= \frac{n!}{k!(n-k)!} P(A_1{}^k \times A_2{}^{n-k}) / P(\Omega^n) \\
&= \frac{n!}{k!(n-k)!} P^n(A_1{}^k \times A_2{}^{n-k}) .
\end{aligned}
$$

On the other hand, H1 together with (3.3) and (3.16) imply that

$$
\begin{aligned}
(3.18) \qquad P^n(A_1{}^k \times A_2{}^{n-k}) &= P^n(S_1 \in A_1,.., S_k \in A_1, S_{k+1} \in A_2,.., S_n \in A_2) \\
&= \prod_{i=1}^{k} P(S_i \in A_1) \prod_{i=k+1}^{n} P(S_i \in A_2)
\end{aligned}
$$

$$= P(S_1 \in A_1)^k \, P(S_1 \in A_2)^{n-k}$$

holds for all $k, n-k \in Z_+$, and similarly that $P^n(A_i{}^n) = P(S_1 \in A_i)^n$ holds for $i=1,2$. Hence, by setting $p_n = P^n(S_1 \in A_1) > 0$ and observing from the definition of $\{A_1, A_2\}$ that $P^n(S_1 \in A_2) = 1 - p_n > 0$, we may conclude that for this choice of $p_n \in (0,1)$,

$$P^n(A_1{}^k \times A_2{}^{n-k}) = p_n{}^k (1-p_n)^{n-k} \tag{3.19}$$

holds identically for all $k \leq n \in Z_+$. Finally, by substituting (3.19) into (3.17), we obtain (3.7), and the result is established. End of Proof.

REMARK: It should be observed that while the relationship between independent population processes and Poisson frequency processes established in Theorem 3.1 represents the most natural statement of this result, the proof of Theorem 3.1 actually implies a somewhat stronger result. In particular, it follows from this proof that the frequency independence hypothesis (H2) need not hold for all finite measurable partitions. For as long as H2 holds for some 2-partition satisfying P2, the total population frequency will be Poisson distributed. Hence by Lemma 3.2, it follows that any measurable k-partition $\{A_1,..,A_k\}$ which does satisfy H2 will continue to yield Poisson distributed frequencies $N_{A1},..,N_{Ak}$ as in (3.6). For example, if dependencies are known to occur only within the subregions of some measurable k-partition $\{A_1,..,A_k\}$, then $N_{A1},..,N_{Ak}$ will have joint distribution (3.6) even though some refinements of the partition $\{A_1,..,A_k\}$ fail to satisfy H2.

3.3.1 Relation to other Formulations of Poisson Processes

To compare the Poisson Frequency Theorem with other approaches to Poisson processes, it is convenient formulate the notion of "rare events" within the present framework of population processes. In their most general form, the two rare-event conditions mentioned in section 2.4.1 above amount to requiring that in any realization of a given population process, P, (i) the chance of any given individual state occurring is zero, and (ii) the chance of two or more identical individual states occurring is also zero. More formally, if for each state $\omega \in \Omega$ we let N_ω denote the frequency variable for the singleton set $\{\omega\} \in M$, then these two requirements correspond, respectively, to the conditions that[14]

(3.20) $$P(N_\omega > 0) = 0 \quad , \qquad \omega \in \Omega$$

(3.21) $$P(\sup_{\omega\in\Omega} N_\omega > 1) = 0$$

Intuitively, these conditions are motivated by the notion of a continuous state space, Ω, in which each individual state has a zero chance of actually occurring. [Indeed, if Ω were a countable set, then condition (i) would exclude all processes except the "null" process defince by $P(\Omega^o) = 1$.] In particular, if the notion of a "continuum" is here taken to be a bounded separable metric space Ω (i.e. possessing a denumerable dense subset), and M is taken to be the Borel σ-algebra generated by the open sets in Ω (with respect to its metric topology), then it can be shown that any population process, P , on (Ω^*, M^*) satisfying conditions (3.20) and (3.21) together with H2 must yield a Poisson frequency process, N , in our sense. [See for example the general results in Kallenberg (1976, Corollary 7.4) and Matthes, et al. (1978, Theorem 1.11.8).][15]

Alternatively, if the frequency process N is postulated to be a function of some volume measure, v , on (Ω, M), then the implicit continuity hypothesis in this case corresponds to the requirement that the measure v be "diffuse" (see section 3.5 below). In this measure-theoretic context, conditions (3.20) and (3.21) are replaced by the corresponding rare-event conditions that[16]

(3.22) $$\lim_{v(A)\to 0} P(N_A > 0) = 0$$

(3.23) $$\lim_{v(A)\to 0} P(N_A > 1 \mid N_A \geq 1) = 0$$

These conditions together with H2 again imply that all such measure-dependent processes must generate Poisson frequency processes (as discussed in section 3.5 below).

Finally, even where such continuity considerations are not appropriate, it is possible to reinterpret the notion of rare events in terms of very "small" individual population processes. With respect to this alternatve view of rare events, one may again characterize Poisson process as superpositions of infinitely many small processes (see section 4.1 below for details).

In summary then, all such characterizations of Poisson processes are seen to involve some notion of rare events. Hence the present approach amounts to replacing this concept with the hypothesis of state independence (H1).[17] Moreover, in all cases in which an explicit population process, P, is postulated, all such rare-event conditions must at least imply an asymptotic form of state independence for P [in the sense of the conditional multinomial property of all Poisson frequencies, as in expression (3.25) below]. Hence it may be argued that state independence is a more general property of Poisson processes in that it continues to be meaningful in cases where the rare-event conditions are not (or are at best very artificial).[18]

3.3.2 State Probability Measures

As an additional consequence of Theorem 3.1, it follows that for each independent population process, P, with mean frequency measure, μ, the associated state probability measure, p, on (Ω, M) given by

$$p(A) = \frac{\mu(A)}{\mu(\Omega)}, \qquad A \in M \tag{3.24}$$

is well defined, and exhibits the following properties:

COROLLARY 3.1 (State Probabilities) *For any independent population process P on (Ω^*, M^*) and measurable k-partition $\{A_1, .., A_k\}$ of Ω,*

$$P^n(N_{A_1} = n_1, \ldots, N_{A_k} = n_k) = n! \prod_{i=1}^{k} \left[\frac{p(A_i)^{n_i}}{n_i!} \right] \tag{3.25}$$

holds for all $n \in Z_+$ and $n_1, .., n_k \in Z_o$ with $\sum_i n_i = n$. Moreover, for any $A \in M$ and $n \in Z_+$,

$$P^n(S_i \in A) = p(A) = \frac{P(A)}{P(\Omega)}, \qquad i = 1,..,n \tag{3.26}$$

Proof: First, (3.25) follows at once from Theorem 3.1 and (3.24) by substituting (3.6) and $P(N_\Omega = n) = [\mu(\Omega)^n / n!] \exp[-\mu(\Omega)]$ into the identity $P^n(N_{A1} = n_1, .., N_{Ak} = n_k) = P(N_{A1} = n_1, .., N_{Ak} = n_k) / P(N_\Omega = n)$. To establish (3.26), observe that if the conditional random variables X_A^i, $i = 1,..,n$, are defined on (Ω^n, M^n, P^n) in terms of the indicator functions in (3.1) by $X_A^i(\omega^n) = \delta_A(\omega_i)$ for all $\omega^n = (.., \omega_i, ..) \in \Omega^n$, then by definition, $E(N_A | N_\Omega = n) = \sum_i E(X_A^i | N_\Omega = n)$ and

$E(X_A^i \mid N_\Omega = n) = P^n(S_i \in A)$. Hence, employing (3.16), we may conclude that

$$(3.27) \qquad E(N_A) = \sum_n E(N_A \mid N_\Omega = n)\, P(N_\Omega = n)$$

$$= \sum_n \Big[\sum_{i=1}^{n} P^n(S_i \in A)\Big]\, P(N_\Omega = n)$$

$$= \sum_n n\, P^n(S_1 \in A)\, P(N_\Omega = n)$$

$$= P^n(S_1 \in A)\, E(N_\Omega)$$

and hence that $P^n(S_i \in A) = P^n(S_1 \in A) = \mu(A)/\mu(\Omega) = p(A)$ for all $i = 1,..,n$. Finally, to establish the second equality in (3.26), observe that for the events $A, \Omega \in M \equiv M^*$, $P(A)/P(\Omega) = P^1(A) = P^1(S_1 \in A) = p(A)$. <u>End of Proof</u>.

<u>REMARK</u>: It is of interest to observe from the proof of this corollary that the existence of state probabilities satisfying (3.26) is guaranteed by H1 alone (together with properties P1 and P2 of all population processes). However, the essential multinomial sampling property of such state probabilities (3.25) depends on Poisson frequencies, and hence implicitly involves H2 as well.

3.4 <u>Independent Flow Processes</u>

As a specific illustration of these results, we now formalize the notion of independent flow processes developed in section 2.5 above. To do so, consider an arbitrary set, X , of flow <u>origins</u>, x , and an arbitrary set, Y , of flow <u>destinations</u>, y . The product space $\Omega = X \times Y$ then denotes the relevant set of all possible <u>flows</u> $\omega = (x,y)$ from X to Y (and is assumed to contain at least two distinct elements). If M_X and M_Y denote fixed σ-algebras of measurable <u>origin zones</u> and <u>destination zones</u> containing the singleton sets in X and Y, respectively, then the relevant family of <u>individual flow events</u> is taken to be the σ-algebra, M , generated by the product set $M_X \times M_Y$. In particular, for any choice of origin zone, X_i , and destination zone, Y_j , the event $\Omega_{ij} = X_i \times Y_j \in M$ denotes the occurrence of a flow from X_i to Y_j . With respect to populations of flows, the product set $\Omega^n = (X \times Y)^n$ contains all possible <u>flow patterns</u>, $\omega^n = ((x_1,y_1),..,(x_n,y_n))$, of size n in Ω . Hence, the relevant set of <u>flow pattern events</u> for populations of size n is taken to be the

σ-algebra M^n generated by all products $A_1 \times \cdots \times A_n$ of individual flow events in M. Finally the population of finite flow patterns is given by $\Omega^* = \cup_{n \in Z_0} \Omega^n$, and the corresponding family of all finite flow pattern events is taken to be the σ-algebra M^* generated by the union $\cup_{n \in Z_0} M^n$.

In this context, all concepts for general population processes have meaningful interpretations in terms of flows. In particular, for any probability measure, P, on (Ω^*, M^*), with associated random flow-frequency variables $N_A: \Omega^* \rightarrow Z_0$ for each measurable set $A \in M$, one may designate P as a flow process whenever conditions P1 and P2 of Definition 3.1 are satisfied. Here, P1 requires that the probability $P(A)$ of any flow event $A \in M^n$, $n \in Z_+$, be invariant with respect to the labeling of flows in each $\omega^n \in A$, and P2 requires that there be at least two disjoint measurable flow events A_1 and A_2 (corresponding possibly to distinct flow origin events $X_1 \times Y$ and $X_2 \times Y$, or to distinct flow destination events $X \times Y_1$ and $X \times Y_2$) for which any flow frequencies N_{A1} and N_{A2} are jointly possible. In particular, P2 implies that any finite flow pattern size, n, is possible, and hence that the conditional flow probabilities $P^n(A) = P(A)/P(\Omega^n)$ are well defined for each $A \in M^n$, $n \in Z_+$. Thus, one may define an associated family of random flow variables, $S_i: \Omega^n \rightarrow \Omega$, $i = 1,..,n$, on (Ω^n, M^n, P^n) with values $S_i(\omega^n) = (x_i, y_i)$ denoting the i-th flow in each possible realized flow pattern $\omega^n = ((x_1,y_1),..,(x_n,y_n))$.

In terms of these flow concepts, one may then designate P as an independent flow process whenever H1 and H2 in Definition 3.2 are also satisfied. In particular, H1 asserts that for any given flow pattern size, n, the individual flows $S_1,..,S_n$ are independently distributed, and H2 asserts that the flow frequencies $N_{A1},..,N_{Ak}$ for any measurable k-partition $\{A_1,..,A_k\}$ of Ω are also independently distributed. For example, the number of flows leaving any disjoint origin zones $X_1,..,X_k$ are independent (where in this case $A_i = X_i \times Y$, $i = 1,..,k$), and similarly, the number of flow terminating at any disjoint destination zones $Y_1,..,Y_k$ are independent (where $A_i = X \times Y_i$, $i = 1,..,k$).

Finally, if these independence conditions can be assumed to hold for a given flow process, then Theorem 3.1 asserts that the family of flow frequencies $(N_A: A \in M)$ must yield a Poisson frequency process, N, on (Ω, M) with associated mean flow-frequency measure, T, defined for all $A \in M$ by $T(A) = E(N_A)$ as in (3.5). Hence the Poisson distributions in (2.9) and associated multinomial distributions given by (3.25) above yield an explicit distribution theory for estimating and testing a variety of mean-frequency models, such as the parametric exponential models illustrated in (2.11) and (2.12) above.

3.5 Homogeneous Poisson Frequency Processes

As a final application of Theorem 3.1, it is of interest to consider the classical notion of homogeneous spatial Poisson processes within the present context. To do so, we first require a notion of "volume" on the relevant state space Ω. In the present setting, the only feature of classical volume (Lebesgue measure on R^n) which is essential for our purposes is its "diffuseness" property that positive volumes of space always be divisible into smaller positive volumes. More precisely, a measure v on (Ω, M) with $0 < v(\Omega) < \infty$ is now designated as a (finite) volume iff for each $A \in M$ with $v(A) > 0$ there is some $B \subset A$ with $0 < v(B) < v(A)$.[19] For example, if the relevant "volume" measure for the shopping-trip flow process on $\Omega = X \times Y$ illustrated in section 2.5.1 above is given by the (measure-theoretic) product of a population measure v_X on M_X and a measure v_Y of shopping opportunities on M_Y, then both v_X and v_Y are required to be diffuse in the sense that all origin zones with positive population and destination zones with positive opportunities are divisible into smaller zones of positive population and opportunities, respectively.

With respect to a given volume measure v on (Ω, M), a frequency process $N = (N_A : A \in M)$ is designated as a homogeneous Poisson frequency process on (Ω, M, v) iff N satisfies (3.6) for all measurable k-partitions of Ω, and in addition, there exists some positive intensity parameter $\lambda > 0$ such that the mean frequency measure, μ, for N satisfies

$$(3.28) \qquad \mu(A) = \lambda v(A)$$

for all $A \in M$. In the classical case, Ω corresponds to a bounded Borel set in R^n (with positive volume), and M and v correspond, respectively, to the Borel subsets of Ω and to the restriction of Lebesgue measure to M.[20]

Finally, if we now designate a population process P on (Ω^*, M^*) as homogeneous with respect to a volume measure v on (Ω, M) iff for all $A_1, A_2 \in M$,

$$(3.29) \qquad v(A_1) = v(A_2) \Rightarrow \mu(A_1) = \mu(A_2)$$

then we obtain the following characterization of homogeneous Poisson frequency processes:

THEOREM 3.2 (Homogeneous Poisson Frequencies) For any given volume measure v on (Ω, M), an independent population process P generates a homogeneous Poisson frequency process N on (Ω, M, v) iff P is homogeneous with respect to v.

Proof: Necessity is immediate since (3.28) implies (3.29). To establish sufficiency, it suffices from Theorem 3.1 to show that (3.28) holds for some $\lambda > 0$. To do so, observe first from the definition of a volume measure v that for each $x \in [0,v(\Omega)]$ there exists some $A \in M$ with $v(A) = x$ [Halmos (1950, Complement 2, Section 42)]. Hence by (3.29), the correspondence, f, defined for all $x \in [0,v(\Omega)]$ by

$$(3.30) \qquad f(x) = \mu(A) \quad \text{for some } A \in M \text{ with } v(A) = x$$

yields a well defined function $f: [0,v(\Omega)] \longrightarrow R_o$ (= nonnegative reals) . Moreover, if for any $x,y \in R_o$ with $x+y \leq v(\Omega)$ we choose sets $A \subseteq B \in M$ with $v(B) = x+y$ and $v(A) = x$, then it follows from the additivity of measures that $v(B-A) = y$, and hence that $f(x+y) = \mu(B) = \mu(A) + \mu(B-A) = f(x) + f(y)$. Thus the nonnegative function f is seen to satisfy Cauchy's equation on the interval $[0,v(\Omega)]$, and we may conclude from the positivity of $f[v(\Omega)] = \mu(\Omega)$ that there must exist some positive constant λ such that for all $x \in [0,v(\Omega)]$ the function f is of the form $f(x) = \lambda x$ [Aczél (1966, Theorem 2.1.4.3)]. Finally, it then follows from (3.30) that for each $A \in M$, $0 \leq v(A) \leq v(\Omega)$ implies that $\lambda v(A) = f[v(A)] = \mu(A)$, and hence that (3.28) must hold. End of Proof.

This axiomatization of homogeneous Poisson frequency processes essentially replaces the rare-event conditions in (3.22) and (3.23) above with the state independence condition (H1), and weakens the usual "homogeneity" condition $v(A_1) = v(A_2) \Rightarrow P(N_{A1}) = P(N_{A2})$ [in axiom (ii) of Karlin and Taylor (1981, Section 16.1)] to condition (3.26) above. Alternatively, one may combine H2 and (3.29) together with the rare-event conditions in (3.20) and (3.21) above to obtain a similar result for the case of a finite volume measure v defined on the Borel σ-algebra M generated by the open sets of a given separable metric space Ω [by employing, for example, Corollary 7.4 in Kallenberg (1976) together with the argument in Theorem 3.2 above]. Finally, it should be noted that while the classical formulation of homogeneous spatial Poisson processes is in terms of the bounded sets in some unbounded metric space Ω , the proof of Theorem 3.2 shows that such processes may also be formulated on bounded spaces Ω [i.e. with finite total volume $v(\Omega)$].

4. APPLICATIONS TO WEAK INTERACTION BEHAVIOR

The basic appeal of the Poisson framework developed here lies in the simplicity and generality of its independence axioms. But such generality has its price. In particular, the nature of "weakly interacting" population behavior is left completely unspecified, and indeed, is only definable implicitly in terms of the absence of certain types of statistical dependencies. Hence in applying this framework to specific population behavior, one is still left with the important practical question of determining whether interactions within this population are "sufficiently weak" to allow its use. Ideally, one would like to have available a general model of population interaction behavior itself, and to establish measurable bounds on the "strength" of these interactions within which the assumptions of statistical independence are appropriate. However, as with the notion of statistical dependence itself, it is perhaps too much to hope for a completely satisfactory general model of interaction behavior. Hence it appears at this time that the only feasible approach to this question is to analyze a wide range of alternative interaction models which are appropriate for certain types of population phenomena, and to examine the robustness of the independence axioms with respect to weak interactions in each of these settings.

With this goal in mind, we shall briefly consider three possible approaches to modeling interaction behavior in this final section which appear to have the widest range of possible applications. In each case, the notion of an independent population process turns out to be the limiting form of these models with respect to an appropriate notion of "weak interactions", and indeed helps to shed some new light on the behavior of these models themselves.

4.1 Superimposed Population Processes

Recall from the discussion of the Weaver-Wilson classification scheme in section 2.4.2 above that one may often view weakly interacting populations as large-scale systems exhibiting disorganized complexity. In this context, a natural model which suggests itself is to treat "disorganized complexity" as the cumulative result of a large number of independent subsystems of behavior. Here the notion of "weak interactions" corresponds to interactions <u>within</u> subsystems, but not <u>between</u> subsystems. For example, in the context of spatial flow processes (sections 2.5 and 3.4 above), each realized pattern of shopping trips or commodity shipments is typically the result of many independent decisions by individual households or firms, each constituting a relevant subsystem within the economy as a whole. Hence, if the travel behavior of each household i (or shipping behavior of each firm i) is modeled as a very "small" flow process P^i, then the resultant spatial flow process P can be taken to be the superposition of all these processes. In particular, if N^i denotes the frequency process generated by P^i, then the overall frequency process N for P can be taken to be the sum $N = \sum_i N^i$ of these individual frequency processes. Hence, one is led to study the asymptotic behavior of the sum of many small frequency processes N^i.

To formalize this notion, we begin by considering a triangular array of frequency processes:

$$(4.1)\qquad \begin{array}{lllll} N^{11}, \ldots, N^{1k_1} & & & & \\ \vdots & \ddots & & & \\ N^{n1}, \ldots, N^{ni}, \ldots, N^{nk_n} & & & & \\ \vdots & & & \ddots & \end{array}$$

over a common measurable space (Ω, M), in which the individual processes N^{ni}, $i = 1,..,k_n$, in each row n are statistically independent. If it is assumed that $k_n \rightarrow \infty$ as $n \rightarrow \infty$, then one may interpret this sequence of rows as approaching an infinite collection $\{N^i\}$ of independent (uniformly) <u>small processes</u>[21] by requiring that the following two conditions hold on each measurable set $A \in M$:

$$(4.2)\qquad \max_{1 \leq i \leq k_n} P(N_A^{ni} > 0) \longrightarrow 0$$

$$(4.3)\qquad \sum_{i=1}^{k_n} P(N_A^{ni} > 1) \longrightarrow 0$$

as $n \rightarrow \infty$. Condition (4.2) requires that the frequencies of processes in each row become uniformly small as $n \rightarrow \infty$, and (4.3) essentially requires that the possibility of any of these frequencies being greater than one can be disregarded as $n \rightarrow \infty$, i.e. that the individual frequencies N_A^{ni} in each set $A \in M$ can eventually be treated as zero-one random variables for large n. In particular, if the cumulative frequency process in each row n is denoted by $N^n = \sum_i N^{ni}$, then (4.3) implies that the associated mean frequencies $\mu_n(A) = E(N_A^n)$ behave like sums of zero-one random variates in the sense that for large n,

$$(4.4)\qquad \mu_n(A) = \sum_{i=1}^{k_n} \mu_i(A) \approx \sum_{i=1}^{k_n} P(N_A^{ni} = 1)$$

In this context, the main property of such infinite superpositions of processes [see for example Çinlar (1972, Theorem 3.10)] is that if their associated mean frequencies converge at all, i.e. if the limits

$$(4.5)\qquad \mu(A) = \lim_{n \rightarrow \infty} \sum_{i=1}^{k_n} P(N_A^{ni} = 1)$$

exist for all $A \in M$, then the sequence of processes N^n converges to the unique Poisson frequency process N with mean frequency measure μ given by (4.5). Hence for large collections of small frequency processes, it may be argued that the resulting superimposed process is approximately Poisson. [In fact, this result is essentially seen to be a generalization of the classical Poisson approximation of the binomial distribution (see for example Feller, 1968)].

To relate this result to the present notion of an independent population process, observe that since each frequency variable N_{Ω}^{ni} is approximately a zero-one variate for large n, it follows that each small process N^i in the limiting sum $N = \sum_i N^i$ contributes at most one point to N.[22] Moreover, since these processes N^i are by definition independent, it is clear that their associated points in Ω will be independently distributed, and hence satisfy state independence (H1) for any given total number of points $N_{\Omega} = n$. Finally since an infinite cumulation $\sum_i N^i$ of such independent processes must surely satisfy frequency independence (H2) as well, it follows that the limiting frequency process can only be generated by an independent population process. Hence the properties of independent population processes help to clarify the intuitive meaning of this convergence to Poisson frequency processes. Conversely, this convergence result shows that whenever population interactions occur only within small independent subpopulations (and are "weak" in this sense), then overall population behavior should correspond closely to an independent population process.

Finally, it is important to emphasize that the above results do not depend on any underlying continuity properties of the state space, and in fact hold for essentially arbitrary spaces [see Remark (3.18) in Çinlar (1972) together with the general formulations of this result in Kallenberg (1976, Corollary 7.5) and Matthes, et al. (1978, Theorem 3.4.4)]. Intuitively, this is made possible by shifting the role of "rare events" from the occurrence of individual states in the continuous case to the occurrence of positive frequencies for each small process in the present case. Indeed, conditions (4.2) and (4.2) are now seen to be the relevant rare-event conditions [as compared with (3.21) and (3.22), or (3.23) and (3.24), respectively].

4.2 Filtered Population Processes

In view of the dependence of superposition processes on an implicit rare-event hypothesis, it is of interest to ask whether it is possible to model weak interactions in a way which involves no notion of rare events. To motivate one possibility here, recall the illustration of weak interaction behavior (section 2.2 above) in which it was asserted that diffuse exposure to colds should not effect the analysis of susceptibility to colds. This assertion can now be made more precise in the following way. Consider first a hypothetical situation in whch all people are continually exposed to cold virus. Under these (very undesirable!) conditions, the realized frequencies of attribute profiles for cold sufferers should indeed be a very good indicator of which attributes are most critical for determining susceptibility to colds. In this context, if the attribute profiles of such cold sufferers were statistically independent (H1), and if frequencies of cold victims were independent across attribute profiles (H2), then one conclude from the Poisson Frequency Theorem that the corresponding mean frequencies could be estimated and tested within the Poisson framework.

However, since all people are not continually exposed to cold virus, one could never observe such a hypothetical process. In reality, those individuals not exposed to colds have effectively been "filtered out" of the potential population of cold victims. Moreover, it is clear that the resulting process can in general be radically different from the hypothetical process above (as in the case of disease X in section 2.2 above). On the other hand, if the degree of exposure is sufficiently diffuse (i.e. if the exposure process itself is in stationary equilibrium), then such effects can often be discounted. In particular, if all individuals have the same chance of being exposed to cold virus, then the resulting "filtered" Poisson process[23] continues to be Poisson. More precisely, if each individual has the same probability, p, of being exposed, then it can be shown [see for example Matthes, et al. (1978, Prop.1.13.7)] that if the hypothetical Poisson process above has mean frequency measure, μ, then the filtered process actually observed will again be Poisson with mean frequency measure, $p\mu$. [In fact, this type of invariance property actually characterizes Poisson processes, as for example in Matthes, et al. (1978, Theorem 10.2.2)]. Hence the Poisson framework can still be employed in this case. Moreover, since the observed means are proportional to the hypothetical means, those attributes critical for susceptibility to colds can still be identified without ever knowing the value of p [and indeed, the resulting state probabilities in (3.24) are seen to be independent of the value of p].

More generally, whenever weak interaction behavior can be treated as a (stationary) filter acting on an independent population process, the resulting population process will again be independent. To see this in an intuitive way, simply think of filters as deleting individual states ω_i from each realized population state $\omega^n = (\omega_1, .., \omega_n)$ with independent and identical probabilities $1 - p$ $(p > 0)$. Then the resulting process is easily seen to satisfy P1, P2, H1, and H2 whenever the original process does. Hence the notion of an independent population process again helps to clarify the characteristic nature of this invariance property of Poisson processes. In addition, it should be clear from simple continuity considerations that if filtering probabilities are approximately uniform then the resulting process will be approximately Poisson. [Explicit results of this type, which relate closely to superposition processes, are given in Westcott (1976) and Matthes, et al. (1978, Prop.10.2.3).]

4.3 Markov Interaction Processes

While filtering does yield an interpretation of weak interactions which requires no notion of rare events, it is still too simplistic to provide a very rich model of interaction behavior itself. All that can be said is that some unobserved process has "deleted" certain potential members from the realized population. With this in mind, we turn finally to a class of Markov processes which provide perhaps the richest interaction models yet developed for analyzing unordered data sets. To motivate the basic idea here, it is instructive to consider a spatial flow example, such as travel behavior within a large urban area. As with the filtering example above, we begin with an idealized noninteractive process, and then introduce the possibility of interactive effects. If we denote the relevant urban area by X, then the state space of potential trips $\omega = (x,y)$ is given by $\Omega = X^2$ (where we now set $Y = X$ in sections 2.5 and 3.4). In this context, we may imagine that potential trips for individual travelers arise independently of one another, and hence define an independent flow process, P_0, over the set M^* of measurable flow events in Ω^* (as in section 3.4). It then follows from the Poisson Frequency Theorem that the frequencies of such potential trips are representable by a Poisson frequency process, N_0, over Ω. Hence, if there were no no interactions among trip makers, one would expect realized trip frequencies to coincide with N_0, and hence to be Poisson distributed.

Now suppose that there are interaction effects among trip makers in terms of traffic congestion. In some situations these interactions may indeed be quite drastic (as evidenced, for example, by the global congestion patterns which can result from a single traffic accident on a major arterial road). However, under "normal" driving conditions, one may assume that such interaction effects are mainly of a local nature, and occur only among trips which are "close together" in some sense. In this context, it is natural to view such normal conditions as a steady-state flow equilibrium in which local interactions among flows lead to no global changes in the system. Models of precisely this type have long been employed in statistical mechanics [see for example Ruelle (1969, sections 2.2 and 3.1) and Kindermann and Snell (1980, sections 1,2, and 3)], and have more recently been extended to a wide range of population phenomena [as for example in Besag (1974), Preston (1975), Strauss (1975), Kelly and Ripley (1976), and Kindermann and Snell (1980, section 6)].

To develop an explicit example, it is convenient to adopt the modeling framework of Kelly and Ripley (1976). We begin by considering any pair of trips $\omega, \omega' \in \Omega$ and write $\delta(\omega,\omega') = 1$ whenever ω interacts with ω', and $\delta(\omega,\omega') = 0$ otherwise. The set $I_\omega = \{\omega' \in \Omega : \delta(\omega,\omega') = 1\}$ then denotes the relevant interactive environment for ω in Ω (which is assumed to be small relative to Ω, as discussed below). In this context, it is then hypothesized that for any current trip pattern $\omega^n = (\omega_1,..,\omega_n)$, the probable entry of trip ω into the system is influenced by the current congestion level, $c(\omega,\omega^n) = \sum_{i=1}^n \delta(\omega,\omega_i)$, for ω in ω^n (i.e. the number of trips currently in the interactive environment of ω), and in particular, is diminished by higher levels of congestion. With respect to this type of "spatial Markov hypothesis", it can be shown that from among those potential trip patterns which arise in P_0, the probable realization of any trip pattern $\omega^n = (\omega_1,..,\omega_n)$ depends only on the total congestion level, $c(\omega^n) = \sum_{i=1}^n c(\omega_i,\omega^n)$, in pattern ω^n. More precisely, it can be shown that (under appropriate regularity conditions) the probability density of such trip patterns with respect to the "reference" process P_0 is always of the form [see Kelly and Ripley (1976, Theorem 1)]:

$$(4.6) \qquad f(\omega^n) = k_\theta \exp[-\theta\, c(\omega^n)]$$

where $\theta \geq 0$ is a constant and k_θ is an appropriate normalizing factor depending on the value of θ. The resulting Markov interaction process, P_θ, defined by the density in (4.6) is thus seen to be "self inhibiting" (for positive θ) in the sense that trip patterns ω^n with higher total congestion levels $c(\omega^n)$ are relatively less likely to occur than they would be in the noninteractive process P_0. Hence, this type of flow process is seen to capture all spatial interaction effects in terms of a single "sensitivity" parameter, θ, which determines the degree to which such interactions influence the process. The detailed properties of this class of Markov interaction processes will be developed in a subsequent paper.

For our present purposes, it suffices to consider the consequences of such processes for weak interaction behavior. We begin by observing that within this richer type of interaction model, it is possible to distinguish two different forms of weak interactions. First of all, interactions may be "weak" in the sense that behavior is not very sensitive to interaction effects, i.e. that θ is close to zero. In this case, it follows at once from the definition of the density in (4.6) that $P_\theta \to P_0$ as $\theta \to 0$, and hence that P_θ is well approximated by the independent flow process P_0 for such weak interactions. On the other hand, interactions may also be "weak" in the sense that congestion levels $c(\omega^n)$ are very small relative to the size of any trip pattern ω^n. Here again it is evident from (4.6) that the resulting flow process should be well approximated by P_0.

This approximation can be made more precise by quantifying the notion of "small" congestion levels. As one possiblity here, the mean frequency measure, μ, for P_0 can be regarded as the relevant "volume" measure for Ω (in the sense of section 3.5 above) which represents the expected volume of potential trips in each region of Ω. Within this context, if the interactive environment I_ω for each potential trip ω is assumed to be small relative to Ω in the sense that the volume $\mu(I_\omega)$ never exceeds some small fraction $p << 1$ of $\mu(\Omega)$, then it can be shown that under these conditions the probable total number of trips in Ω for process P_θ is approximated by

$$P_\theta(N_\Omega = n) \approx \frac{\mu(\Omega)^n}{n!}\exp[-\mu(\Omega)]\left\{1 - p\theta\,\frac{n(n-1)}{2}\right\} \tag{4.7}$$

Hence if p is very small, then the trip frequencies in (4.7) are indeed seen to be approximately Poisson distributed -- except for very large n. In other words, for p sufficiently small, the Poisson approximation can be expected to hold for all trip patterns except those far above the expected volume $\mu(\Omega)$ of potential trips in Ω. Moreover, it is clear from (4.7) that if θ is also very small, then these two types of weak interaction effects work multiplicatively to yield even better approximations. A more detailed analysis of such approximation results will be presented in a subsequent paper.

5. SELECTED REFERENCES

Aczél, J. (1966) Lectures on Functional Equations and their Applications, Academic Press, New York.

Bhattacharyya, G.K. and R.A.Johnson (1977) Statistical Concepts and Methods, Wiley, New York.

Besag, J.E. (1974) "Spatial Interaction and the Statistical Analysis of Lattice Systems", Journal of the Royal Statistical Society, Series B, Vol.36, pp.192-236.

Bishop, Y.M.M., S.E.Fienberg, and P.W.Holland (1975) Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge, Mass.

Carter, D.S. and P.M.Prenter (1972) "Exponential Spaces and Counting Processes", Z.Wahrscheinlichkeitstheorie, Vol. 21, pp. 1-19.

Chatterji, S.D. (1963) "Some Elementary Characterizations of the Poisson Distribution", American Mathematical Monthly, Vol. 70, pp. 958-964.

Çinlar, E. (1972) "Superposition of Point Processes", in Lewis, P.A. (ed.) Stochastic Point Processes: Statistical Analysis, Theory, and Applications, Wiley, New York, pp. 549-606.

Cox, D.R. and V.Isham (1980) Point Processes, Chapman and Hall, London.

Daley, D.J. and D.Vere-Jones (1972) "A Summary of the Theory of Point Processes", in Lewis, P.A. (ed.) Stochastic Point Processes: Statistical Analysis, Theory, and Applications, Wiley, New York, pp. 299-383.

Feller, W. (1968) An Introduction to Probability Theory and its Applications, Vol.1, Wiley, New York.

Flowerdew, R. and M.Aitkin (1982) "A Method of Fitting the Gravity Model based on the Poisson Distribution", Journal of Regional Science, Vol.22, pp. 191-202.

Gray, R.H. and A.K.Sen (1983) "Estimating Gravity Model Parameters: A Simplified Approach based on the Odds Ratio", Transportation Research, Vol.17B, pp. 117-132

Haberman, S.J. (1964) The Analysis of Frequency Data, University of Chicago Press, Chicago.

Haberman, S.J. (1979) Analysis of Qualitative Data (two volumes) Academic Press, New York.

Haight, F.A. (1967) Handbook of the Poisson Distribution, Wiley, New York.

Halmos, P.R. (1950) Measure Theory, Von Nostrand, Princeton, New Jersey.

Johnson, N.L. and S.Kotz (1969) Discrete Distributions, Houghton Mifflin Co., Boston.

Kagan, A.M., Y.V.Linnik, and C.R.Rao (1973) Characterization Problems in Mathematical Statistics, Wiley, New York.

Kallenberg, O. (1976) Random Measures, Academic Press, New York.

Karlin, S. and H.M.Taylor (1981) A Second Course in Stochastic Processes, Academic Press, New York.

Katz, A. (1967) Principles of Statistical Mechanics, Freeman and Co., San Francisco.

Kelly, F.P. and B.D.Ripley (1976) "A Note on Strauss' Model for Clustering", Biometrika, Vol. 63, pp. 357-360.

Kindermann, R. and J.L.Snell (1980) Markov Random Fields and their Applications, Contemporary Mathematics Series, Vol. 1, American Mathematical Society, Providence, Rhode Island.

Lukacs, E. (1970) Characteristic Functions, Griffin, London.

Matthes, K., J.Kerstan, and J.Mecke (1978) Infinitely Divisible Point Processes, Wiley, New York.

Moyal, J.E. (1962) "The General Theory of Stochastic Population Processes", Acta Mathematica, Vol.108, pp. 1-31.

Nishisato, S. (1980) Analysis of Categorical Data: Dual Scaling and its Application, University of Toronto Press, Toronto, Canada.

Patel, J.K., C.H.Kapadia, and D.B.Owen (1976) Handbook of Statistical Distributions, Marcel Dekker Inc., New York.

Patil, G.P. and V.Seshadri (1964) "Characterization Theorems for some Univariate Probability Distributions", Journal of the Royal Statistical Society, Series B, Vol. 26, pp. 286-292.

Plackett, R.L. (1974) The Analysis of Categorical Data, Griffin, London.

Preston, C.J. (1975) "Spatial Birth-and-Death Processes", Bulletin of the International Statistical Institute, Vol. 46, pp. 371-391.

Raikov, D. (1938) "On the Decomposition of Gauss' and Poisson's Laws", Izvestia Akademie Nauk SSSR, Series A, pp.91-124 (in Russian).

Ruelle, D. (1969) Statistical Mechanics, Benjamin, New York.

Sen, A.K. and S.Sööt (1981) "Selected Procedures for Calibrating the Generalized Gravity Model", Papers of the Regional Science Association, Vol. 48 , pp. 165-176.

Snyder, D.L. (1975) Random Point Processes, Wiley, New York.

Strauss, D.J. (1975) "A Model for Clustering", Biometrika, Vol. 62, pp. 467-475.

Tjur, T. (1980) Probability based on Radon Measures, Wiley, New York.

Volodin, I.N. (1965) "On Distinguishing between Poisson and Polya Distributions on the basis of a Large Number of Samples", Teoriya Veroyatnostei i ee Primeneniya, Vol. 10, pp. 364-367 (in Russian).

Weaver, W. (1958) "A Quarter Century in the Natural Sciences", Annual Report, Rockefeller Foundation, New York.

Westcott, M. (1976) "Simple Proof of a Result on Thinned Point Processes", Annals of Probability, Vol. 4, pp. 89-90.

Wilson, A.G. (1974) "Urban and Regional Models: Some Organising Concepts to Describe the State of the Art", in Baxter, R., M.Echenique, and J.Owers (eds.) Urban Development Models, MTP, Lancaster, England.

Wilson, A.G. (1977) "Research Methods for Urban and Regional Modelling" in Wilson, A.G., P.H.Rees, and C.M.Leigh (eds.) Models of Cities and Regions: Theoretical and Empirical Developments, Wiley, New York.

FOOTNOTES TO TEXT

(1) Note in particular that this condition implies that the potential size of the given population can be unboundedly large. Hence it is implicitly assumed here that the actual physical bounds on population size are much larger than those values typically observed in practice. At the other extreme, this condition also implies that a zero population size is always possible, as for example, when a disease is completely eradicated or a political party is dissolved.

(2) Even with noncontagious diseases, however, there may continue to be important indirect interaction effects. In diseases such as lung cancer, for example, if smoking contributes to the disease, and if friends of smokers tend to be smokers, then there may exist significant attribute dependencies among victims of the disease.

(3) As in the case of the positive divisibility condition (P2), this hypothesis can never hold exactly for bounded populations, since knowledge of the number of members occupying any state automatically restricts the numbers which can occupy other states. Hence, as in footnote (1) above, it is implicitly assumed that the maximum size of the population is sufficiently large to be ignored for all practical purposes.

(4) It is also important to note here that even if the disease occurs exclusively among males, the number of male victims does not influence the number of female victims -- which is identically zero.

(5) As observed by Plackett (1974) and others, the positivity of mean frequencies leads quite naturally to the use of log linear rather than linear models.

(6) An alternative formal description of such a sampling procedure is given by Tjur (1980, Theorem 11.4.5) in which it is shown that if the cumulative mean frequencies for a given sequence of independently and identically distributed random variables converge at all, then the limiting frequency distribution must be Poisson.

(7) It is of interest to observe here that the original idea of "disorganized complexity" proposed by Weaver (1958) was largely inspired by precisely this type of equilibrium concept as employed in statistical mechanics. [See for example the discussion of such steady state equilibria in Katz (1967, Chapter 6) and Ruelle (1969, Section 1.4)].

(8) There are of course many situations in which this assumption may fail to hold. For example, mass media advertising can lead to changes in the purchase decisions of large numbers of households. Similarly, news reports of job opportunities can lead to large scale migration flows from areas of unemployment. Hence in the present analysis it is assumed that the relevant system of flows is not undergoing such global changes.

(9) Separation measures such as travel time and travel cost between zones i and j are implicitly assumed to be average values over the flows in Ω_{ij}. Alternatively, such measures may correspond to the specific travel time or cost between representative "central locations" $x_i \in X_i$ and $y_j \in Y_j$, respectively.

(10) The process $N = (N_A: A \in M)$ is designated as a <u>counting process</u> by Moyal (1962), and is alternatively designated as a <u>point process</u> by many authors. [See for example Kallenberg (1976).]

(11) Our definition of a population process specializes the more general definition in Moyal (1962) [which is also designated as a finite point process by Daley and Vere-Jones (1972) and others]. In particular, P1 corresponds to Moyal's condition that a population process be symmetric. Our positive divisibility condition P2 appears to have no counterpart in Moyal's framework (and indeed corresponds more closely to the diffuseness condition on volume measures employed in the axiomatization of homogeneous spatial point processes, as developed in section 3.5 below). In addition, it should be noted that the interchangeability condition P1 can be incorporated directly into the definition of the population state space by employing unordered sequences rather than ordered sequences. [See for example the elegant formulation of this approach in terms of "exponential sets" in Carter and Prenter (1972).]

(12) For since $\{A_1, A_2\}$ partitions Ω, it follows by setting $n_1 = n$ and $n_2 = 0$ in P2, we must have $P(N_\Omega = n) \geq P(N_{A1} = n, N_{A2} = 0) > 0$.

(13) This measure is also called the intensity measure for the point process defined by the frequency process $N = (N_A : A \in M)$, as for example in Matthes, et al. (1978, Section 1.2).

(14) Conditiōn (3.20) corresponds to the requirement in Kallenberg (1976, p.45) that the random frequency measure N have no fixed atoms, and equivalently, to the condition in Matthes, et al. (1978, Section 1.1) that the corresponding probability distribution P_N of the random frequency measure N be continuous. Condition (3.21) corresponds to the requirement in Matthes, et al. (1978, Section 4.1) that P_N assign probability one to the family of simple realizations of N, and equivalently, to the requirement in Kallenberg (1976, p.14) that P_N be almost surely simple.

(15) Frequency independence (H2) corresponds to the independent increments condition in Kallenberg (1976, p.8), and to the condition in Matthes, et al. (1978, p.16) that the distribution of the frequency process N be free from after-effects.

(16) Conditions (3.22) and (3.23) correspond to the second part of axiom (ii) and to axiom (iv), respectively, in Karlin and Taylor (1981, Section 16.1).

(17) Other characterizations of the Poisson distribution are possible which do not involve rare events, including the behavior of its moment-generating function and sample statistics. See for example the many results summarized in Haight (1967, Chapter 2), Johnson and Kotz (1969, Section 4.5), and Patel, et al. (1976, Section 5.2.2).

(18) However, from a formal viewpoint it should be emphasized that while the stat independence condition requires a specification of the basic population process P ((Ω^*, M^*), the rare-events conditions do not. In particular, all of these conditi can be formulated directly in terms of frequency processes, N [i.e. integer-valued random measures, in the terminology of Kallenberg (1976)], and in this sense, can b said to be meaningful even when no explicit population process is defined.

(19) Such diffuse measures are also designated as <u>nonatomic</u> [Halmos (1950, p.174)]

(20) The restriction to a bounded subset Ω of R^n ensures that $v(\Omega) < \infty$, and hen that $\mu(\Omega) = \lambda v(\Omega) < \infty$, as required by our definition of a Poisson frequency process.

(21) Such processes are also designated as <u>sparse processes</u>, as for example in Cox and Isham (1980).

(22) More precisely, each small process N^i in the limiting sum $\sum_i N^i$ must correspond to a zero-one random measure on M satisfying $N^i_A \in \{0,1\}$ for all $A \in M$. Hence, e realization of N^i must be either an indicator function δ_ω (i.e. Dirac measure) for some $\omega \in \Omega$ (representing the realization of individual state ω) or must be the nu measure on M (representing the occurrence of no realization in Ω). [See Kallenber (1976, Section 1) for further discussion of such random measures.]

(23) For a detailed treatment of such processes, see Snyder (1975, Chapter 4). Not also that filtered processes are equivalently designated as <u>thinned processes</u> or <u>censored processes</u> by many authors.