THE DETECTION AND ESTIMATION OF TRENDS IN TWO-DIMENSIONAL POISSON PROCESSES

Robert Dale Rantschler

Naval Postgraduate School
Monterey, California

March 1973

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

THE DETECTION AND ESTIMATION OF

TRENDS IN TWO-DIMENSIONAL POISSON PROCESSES

by

Robert Dale Rantschler

Thesis Advisor: P.A.W. Lewis

March 1973

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3 REPORT TITLE

The Detection and Estimation of Trends in Two-Dimensional Poisson Processes

4 DESCRIPTIVE NOTES (Type of report and, inclusive dates)

Master's Thesis; March 1973

5 AUTHOR(S) (First name, middle initial, last name)

Robert Dale Rantschler

6 REPORT DATE March 1973	7a. TOTAL NO. OF PAGES 71	7b. NO. OF REFS 17
8a. CONTRACT OR GRANT NO b. PROJECT NO c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	

10 DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

The occurrence of independent events at random in the plane, i.e. the formation of a planar point process, is discussed. Both homogeneous and nonhomogeneous processes are considered. A specific functional form for the parameter in a nonhomogeneous planar Poisson process is used to illustrate the development of test and parameter estimation techniques. The problem finds application in the desription of biological phenomena as well as in search and detection problems.

DD FORM 1 NOV 68 1473 (PAGE 1)

S/N 0101-807-6811

Security Classification

A-31408

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
point process						
stochastic process						
planar process						
poisson process						
bivariate poisson process						
non-homogeneous process						
parameter detection						
parameter estimation						
maximum likelihood estimation						
spatial distributions						
two dimensional process						
conditional likelihood						
uniform conditional test						
Durbin's modification						
distance methods						
rank correlation methods						
point estimation						
confidence interval estimation						
significance level						

The Detection and Estimation of Trends in Two-Dimensional Poisson Processes

by

Robert Dale Rantschler
Lieutenant, United States Navy
B.S.(E.E.), University of Colorado, 1964

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1973

Author Robert D Rantschler

Approved by: Peter A.W. Lewis
Thesis Advisor

Chairman, Department of Operations Research
and Administrative Sciences

Academic Dean

ABSTRACT

The occurrence of independent events at random in the plane, i.e. the formation of a planar point process, is discussed. Both homogeneous and nonhomogeneous processes are considered. A specific functional form for the parameter in a nonhomogeneous planar Poisson process is used to illustrate the development of test and parameter estimation techniques. The problem finds application in the description of biological phenomena as well as in search and detection problems.

TABLE OF CONTENTS

I. INTRODUCTION

Many problems arising naturally in a physical sense are often so complex that the identification and description of underlying mechanisms must use the tools of probability and statistics. Some of the reasons leading to the requirement of using these tools are:

(i) the data base may be so large or complex as to preclude identification of any driving mechanism without recourse to statistical analysis;

(ii) if identifiable, the mechanisms may be inherently probabilistic; or

(iii) if identifiable and deterministic, the governing law which the mechanisms obey may be unknown.

This paper is concerned with the use of statistics in the identification and mathematical description of the spatial distribution of events (occurrences). Included is the detection and estimation of parameters which influence the description of this distribution.

The area of concern here is a departure from those statistical methods which have been developed to detect the effect of varying a controlled segment of the underlying mechanism. Among those methods would be the design of experiments, regression analysis, time series analysis, and analysis of variance. One goal of such analysis is to hopefully predict the advisibility of pursuing some course of action.

In the basic model of this paper, events are considered to occur with a Poisson distribution in the plane. This "is the natural model for the expression that 'points are distributed at random'," [Fisher, 1972, p. 141]. The bivariate Poisson process will be defined and then developed through the use of partial differential-difference equations, a widely repeated procedure in the univariate case but neglected in the bivariate case.

Initially a homogeneous Poisson process will be assumed to control the underlying mechanisms. Then trends will be introduced by defining the Poisson parameter in such a way as to make it be spatially dependent. This will be the basis for the definition of the non-homogeneous Poisson process. Time inhomogeneity will not be considered. Thus, the data are assumed to be taken concurrently, i.e., the period of observation is short compared to any period of change of the parameters.

Tests will be developed to distinguish between homogeneity and non-homogeneity and the method of maximum likelihood will be used to develop estimates of the parameter in the homogeneous case and parameters in the non-homogeneous case. In the latter case, conditional likelihood techniques will be utilized to develop tests and estimates. Throughout, testing and estimation procedures will be based on a single realization of the process which consists of the number of events observed and their spatial locations.

The problem of concern finds application in the estimation of the density of trees in a forest; here one might be concerned with estimating the potential yield of lumber from a given forest area where inhomogeneities arise due to soil, weather patterns, topography and other physical reasons.

Another application might be in naval search and detection problems. For example, one might be searching for a merchant ship in distress whose location is not known exactly due to failure of the ship's communication equipment. Here the independence assumptions of the planar Poisson process may be valid, but not the assumption of homogeneity. Inhomogeneities of location occur because of preferred sea lanes and physical characteristics of the ocean and atmosphere.

II. THE HOMOGENEOUS POISSON PROCESS IN THE PLANE (HPPP)

A. GENERAL DEVELOPMENT

Consider a stochastic process of events occurring in the plane (i.e., a so-called point process) which is characterized by the assumptions

I. There exists a finite positive constant $\lambda > 0$.

II. For any integer $k \geq 1$ and any set of non-overlapping regions $R_1, \cdots, R_k$ with areas $A_1, \cdots, A_k$, (in the usual geometric sense), the number of events occurring in any region i, denoted $N(R_i)$, has a Poisson distribution with parameter λA_i which depends only on the area of the region, A_i, and not its shape. Thus,

$$\text{prob}\{N(R_i) = n_i\} = \frac{(\lambda A_i)^{n_i} \exp(-\lambda A_i)}{n_i!} . \quad (1)$$

III. Further, $N(R_i)$, $i = 1,2,\cdots,k$, are mutually independent in that $N(R_i)$ is not affected by the occurrence of events in any other region or in any grouping of the regions, G, as long as $R_j \cap G = \emptyset$. Thus

$$\text{prob}\{N(R_i)=n_i,\ i=1,\cdots,k\} = \prod_{i=1}^{k} \frac{(\lambda A_i)^{n_i} e^{-\lambda A_i}}{n_i!} \quad (2)$$

Definition 1: If a process obeys the above assumptions it is called a homogeneous planar Poisson process (HPPP).

For reasons of arbitrary shape the above basic definitions will suffice. However, under certain geometrical assumptions, an

equivalent definition for the HPPP can be achieved in a manner similar to the development of the univariate Poisson process through the use of partial differential-difference equations. This is useful for the development of statistical properties and will be very important in the development of the non-homogeneous process. Such a development also provides another phenomenological approach to the homogeneous Poisson process, one which might arise through the structuring of a model for instance. For illustrative purposes the following development will be accomplished using rectangular regions. Note that the development is very dependent on the geometry involved; hence developments with other geometries (e.g. circular regions) must proceed somewhat differently.

The underlying assumptions in the differential equation development will be

I'. There exists a finite positive constant $\lambda > 0$.

II'. For any region R* with incremental area ΔA, independent of the shape of the region except possibly as noted above

(a) prob {no event in R*} = $1 - \lambda\Delta A + o(\Delta A)$,

(b) prob {one event in R*} = $\lambda\Delta A + o(\Delta A)$,

(c) prob {more than one event in R*} = $o(\Delta A)$,

where "$g(\Delta A)$ is $o(\Delta A)$" means $\lim_{\Delta A \to 0} \frac{g(\Delta A)}{\Delta A} = 0$, or specifically in rectangular regions the limit <u>as Δx or Δy or both</u> go to zero of $\frac{g(\Delta A)}{\Delta x \Delta y}$ is zero.

III'. The occurrence of events in R^* is independent of the occurrence of events in any region R^+ where $R^* \cap R^+ = \emptyset$.

It will be shown that I', II' and III' imply and are implied by I, II and III so that the two sets of assumptions are equivalent and hence the incremental assumptions give rise to a HPPP. Clearly I and I' are the same, as are III and III'. Also II implies II' since by (1)

$$\text{(a)} \quad \text{prob}\{N(R^*) = 0\} = e^{-\lambda\Delta A} = 1 - \lambda\Delta A + \frac{\lambda^2}{2}(\Delta A)^2 - \ldots$$

$$= 1 - \lambda\Delta x\Delta y + \frac{\lambda^2}{2}\Delta x^2\Delta y^2 - \ldots$$

$$= 1 - \lambda\Delta A + o(\Delta A),$$

with the definition of $o(\Delta A)$ given above. Also

$$\text{(b)} \quad \text{prob}\{N(R^*) = 1\} = \lambda\Delta A e^{-\lambda\Delta A} = \lambda\Delta A(1 - \lambda\Delta A + \ldots)$$

$$= \lambda\Delta A + o(\Delta A)$$

and $$\text{(c)} \quad \text{prob}\{N(R^*) \geq 2\} = \sum_{i=2}^{\infty} \frac{(\lambda\Delta A)^i e^{-\lambda\Delta A}}{i!} = o(\Delta A).$$

The problem remaining in order to demonstrate equivalence between the two sets of assumptions is to show that II' implies II.

Consider a region R bounded by the co-ordinate axes and lines $x = X^*$ and $y = Y^*$, with area X^*Y^*. Now extend the

sides to $x = X^*+\Delta x$ and $y = Y^*+\Delta y$ (see Figure 1). Consider the probability of n events occurring in the extended region, $R' = R + R_1 + R_2 + R_3$, where:

(a) R has area X^*Y^*,

(b) R_1 has area $X^*\Delta y$,

(c) R_2 has area $Y^*\Delta x$,

(d) R_3 has area $\Delta x\Delta y$;

(a)-(d) imply R' has area $X^*Y^* + X^*\Delta y + Y^*\Delta x + \Delta x\Delta y$.

The assumptions I', II', and III' imply

$$\begin{aligned} &\text{prob }\{\text{no event in } R_1\} = 1 - \lambda X^*\Delta y + o(X^*\Delta y),\\ &\text{prob }\{\text{one event in } R_1\} = \lambda X^*\Delta y + o(X^*\Delta y), \qquad (3)\\ &\text{prob }\{\text{more than one event in } R_1\} = o(X^*\Delta y); \end{aligned}$$

$$\begin{aligned} &\text{prob }\{\text{no event in } R_2\} = 1 - \lambda Y^*\Delta x + o(Y^*\Delta x),\\ &\text{prob }\{\text{one event in } R_2\} = \lambda Y^*\Delta x + o(Y^*\Delta x), \qquad (4)\\ &\text{prob }\{\text{more than one event in } R_2\} = o(Y^*\Delta x); \end{aligned}$$

and

$$\begin{aligned} &\text{prob }\{\text{no event in } R_3\} = 1 - \lambda\Delta x\Delta y + o(\Delta x\Delta y),\\ &\text{prob }\{\text{one event in } R_3\} = \lambda\Delta x\Delta y + o(\Delta x\Delta y), \qquad (5)\\ &\text{prob }\{\text{more than one event in } R_3\} = o(\Delta x\Delta y). \end{aligned}$$

Moreover, statements (3), (4), and (5) are independent.

It is noted that the above equations may have two different interpretations. For instance in (3), prob{one event in R_1} = $X^*\Delta y + o(X^*\Delta y)$ is interpreted to mean one event in a two-dimensional process with parameter λ and area $X^*\Delta y$.

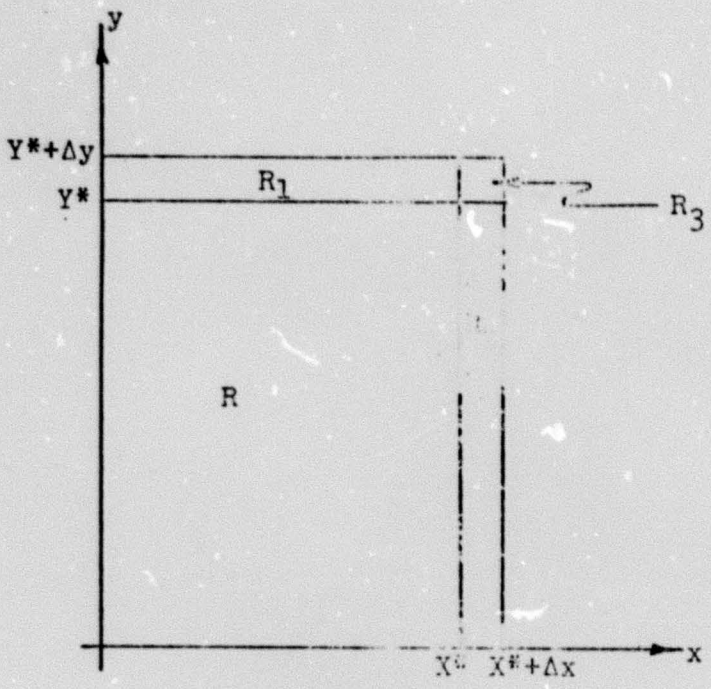

Figure 1. The incremental increase of a rectangular region.

However, another interpretation would be to consider the one-dimensional (marginal) process of events projected onto the y-axis, in which case the parameter is λX^* and the incremental interval has length Δy.

For notational convenience, let $P_n(X^*,Y^*)$ denote the probability that n events occur in a region with area X^*Y^*. The differential-difference equations are written noting that n events may occur in an extended region by having n events in the unextended region and no events in the extension, n-1 events in the unextended region and one event in the extension, etc. Hence

$$\begin{aligned} P_n(X^*+\Delta x,Y^*) &= P_n(X^*,Y^*) \cdot P_o(\Delta x,Y^*) \\ &\quad + P_{n-1}(X^*,Y^*) \cdot P_1(\Delta x,Y^*) \\ &\quad + P_{n-2}(X^*,Y^*) \cdot P_2(\Delta x,Y^*) + \ldots \\ &= P_n(X^*,Y^*)[1-\lambda Y^*\Delta x] + P_{n-1}(X^*Y^*)[\lambda Y^*\Delta x] + o(Y^*\Delta x). \end{aligned} \tag{6}$$

Similarly,

$$\begin{aligned} P_n(X^*,Y^*+\Delta y) &= P_n(X^*,Y^*)[1-\lambda X^*\Delta y] + P_{n-1}(X^*,Y^*)[\lambda X^*\Delta y] + \\ &\quad + o(X^*\Delta y), \end{aligned} \tag{7}$$

and

$$P_n(X^*+\Delta x, Y^*+\Delta y) = P_n(X^*,Y^*)[1-\lambda Y^*\Delta x][1-\lambda Y^*\Delta y][1-\lambda\Delta x\Delta y]$$

$$+ P_{n-1}(X^*,Y^*)[\lambda Y^*\Delta x(1-\lambda X^*\Delta y)(1-\lambda\Delta x\Delta y)$$

$$+ \lambda X^*\Delta y(1-\lambda Y^*\Delta x)(1-\lambda\Delta x\Delta y)$$

$$+ \lambda\Delta x\Delta y(1-\lambda X^*\Delta y)(1-\lambda Y^*\Delta x)]$$

$$+ P_{n-2}(X^*,Y^*)[\lambda X^*\Delta y \cdot \lambda Y^*\Delta x(1-\lambda\Delta x\Delta y)$$

$$+ \lambda X^*\Delta y(1-\lambda Y^*\Delta x)\lambda\Delta x\Delta y$$

$$+ (1-\lambda X^*\Delta y)\lambda Y^*\Delta x\lambda\Delta x\Delta y] \qquad (8)$$

$$+ P_{n-3}(X^*,Y^*)[\lambda^3 X^*Y^*\Delta x^2\Delta y^2]$$

$$+ o(\Delta x\Delta y) + o(X^*\Delta y) + o(Y^*\Delta x).$$

Interpreting the above equations, the third term on the right hand side of (8), for example, states that there can be n events in the extended region R' if there are n-2 events in R and exactly one event in each of any two of the added regions. That is, there can be two events in the added regions R_1, R_2 and R_3 if one occurs in each of two regions and none occurs in the third region, i.e., one in R_1, one in R_2 and none in R_3, etc. Collecting all terms of order $o(\Delta x\Delta y)$, $o(X^*\Delta y)$ and $o(Y^*\Delta x)$, (8) reduces to

$$P_n(X^*+\Delta x,Y^*+\Delta y)=P_n(X^*,Y^*)[1-\lambda Y^*\Delta x-\lambda X^*\Delta y-\lambda\Delta x\Delta y+\lambda^2X^*Y^*\Delta x\Delta y]$$

$$+ P_{n-1}(X^*,Y^*)[\lambda Y^*\Delta x-2\lambda^2X^*Y^*\Delta x\Delta y+\lambda X^*\Delta y+\lambda\Delta x\Delta y]$$

$$+ P_{n-2}(X^*,Y^*)[\lambda^2X^*Y^*\Delta x\Delta y]+o(\Delta x\Delta y)+o(X^*\Delta y)+o(Y^*\Delta x). \quad (8')$$

$$= P_n(X^*,Y^*)[1-\lambda Y^*\Delta x]+P_{n-1}(X^*,Y^*)[\lambda Y^*\Delta x]+o(Y^*\Delta x)$$

$$+ P_n(X^*,Y^*)[1-\lambda X^*\Delta y]+P_{n-1}(X^*,Y^*)[\lambda X^*\Delta y]+o(X^*\Delta y)$$

$$- P_n(X^*,Y^*)-\lambda P_n(X^*,Y^*)\Delta x\Delta y+\lambda P_{n-1}(X^*,Y^*)\Delta x\Delta y$$

$$+ \lambda^2X^*Y^*[P_n(X^*,Y^*)-2P_{n-1}(X^*,Y^*)+P_{n-2}(X^*,Y^*)]\Delta x\Delta y+o(\Delta x\Delta y).$$

Noting from equations (6) and (7) that the first three terms on the right-hand side of the above equation are $P_n(X^*+\Delta x,Y^*)$ and the next three are $P_n(X^*,Y^*+\Delta y)$ and rewriting (8'), the result is

$$P_n(X^*+\Delta x,Y^*+\Delta y)=P_n(X^*+\Delta x,Y^*)+P_n(X^*,Y^*+\Delta y)-P_n(X^*,Y^*)$$

$$- \lambda P_n(X^*,Y^*)\Delta x\Delta y+\lambda P_{n-1}(X^*,Y^*)\Delta x\Delta y \quad (8'')$$

$$+ \lambda^2X^*Y^*[P_n(X^*,Y^*)-2P_{n-1}(X^*,Y^*)+P_{n-2}(X^*,Y^*]\Delta x\Delta y$$

$$+ o(\Delta x\Delta y).$$

The definition of the second partial derivative with respect to two variables is

$$\frac{\partial}{\partial y}\left(\frac{\partial F(x,y)}{\partial x}\right)=\lim_{\Delta y\to 0}\frac{1}{\Delta y}\left\{\lim_{\Delta x\to 0}\frac{F(x+\Delta x,y+\Delta y)-F(x,y+\Delta y)}{\Delta x}\right.$$

$$\left.-\lim_{\Delta x\to 0}\frac{F(x+\Delta x,y)-F(x,y)}{\Delta x}\right\}.$$

Hence, transposing the first three terms of equation (8") to the right hand side, dividing by $\Delta x\Delta y$ and then taking the double limit results in

$$\frac{\partial^2 P_n(X^*,Y^*)}{\partial x\partial y}=-\lambda P_n(X^*,Y^*)+\lambda P_{n-1}(X^*,Y^*)+\lambda^2 X^*Y^*[P_n(X^*,Y^*)$$

$$-2P_{n-1}(X^*,Y^*)+P_{n-2}(X^*,Y^*)] \qquad (9)$$

The solution to (9) (a partial differential-difference equation) is easily shown to be

$$P_n(X^*,Y^*)=\frac{K(\lambda X^*Y^*)^n\exp(-\lambda X^*Y^*)}{n!}\ ,\quad n=2,3,\cdots \qquad (10a)$$

where K is an arbitrary constant. Special considerations are needed for n = 0,1 since for these cases some of the terms in (8") and (9) are not defined. Rewriting (8") and (9) while concurrently eliminating the proper terms leads to

$$P_n(X^*,Y^*)=\frac{K(\lambda X^*Y^*)^n\exp(-\lambda X^*Y^*)}{n!}\ ,\quad n=0,1,2,\ldots \qquad (10b)$$

Since $P_n(X^*,Y^*)$ is a probability statement and for any given region the number of events in that region must be some non-negative integer, the constant K is seen to be unity.

Hence (10b) is equivalent to (1) which was to be proven. Thus the two sets of assumptions imply the same things, namely that the number of events in a region has a Poisson distribution with parameter proportional to the area of the region and independent of its shape and the number and position of events outside the region. Note that the formulation excludes multiple events, i.e., the occurrence of two or more events at any point or on any line segment added region such as R_1 in Figure 1.

Also a similar derivation will go through for circular regions using polar co-ordinates, but there are differences in the special properties of the Poisson process as defined through assumptions I, II and III in differently shaped regions. These are discussed below. The differences in the special properties of the non-homogeneous planar Poisson processes as they vary with different geometries is an essential element of the analysis of points (events) in the plane.

B. TESTING DATA FOR HOMOGENEOUS PLANAR POISSON PROCESS (HPPP)

Given the occurrence and spatial location of n events in a rectangular region of area X*Y*, consider the problem of determining whether or not these points occur as realizations of the HPPP. Miles [1970,p.89] has stated a consequence of definition 1 as

Corollary. Assume a rectangular region R_i with area A_i. Given $N(R_i) = n$ and $0 < A_i < \infty$, the n points are independently and uniformly distributed in R_i.

Proof: Let $A_1 = X^*Y^*$ where R_1 is a rectangular region bounded by the coordinate axes, $x = X^*$ and $y = Y^*$. Label the n given points in any convenient manner, e.g., on the magnitude of the y-component. Let $(x,y)_{(i)}$ denote the i^{th} labelled event. Consider an incremental region with area dxdy which has the property that $\text{prob}\{\text{exactly one event in the incremental region of } dxdy\} = P_1(dx,dy) = \lambda dxdy + o(dxdy)$. Consider now n incremental rectangles $dx_i dy_i$, $i = 1,\cdots,n$, placed in R_1. Ignoring probabilities of $o(dx_i dy_i)$, assumptions I, II and III imply that the joint probability that the i^{th} event falls in the incremental rectangle, $dx_i dy_i$, $i = 1,\cdots,n$, and exactly n events occur altogether in X^*Y^* is given by

$$\lambda dx_1 dy_1 \cdots \lambda dx_n dy_n \exp\{-\lambda X^* Y^*\}.$$

Restating in terms of the density function,

$$f\{(x,y)_{(1)},\ldots,(x,y)_{(n)},n;\lambda\} = \lambda^n \exp\{-\lambda X^* Y^*\},$$

where $f\{\ldots\}$ is the joint density of $(x,y)_{(i)}$, $i = 1,\ldots,n$, and the probability that the number of events in X^*Y^* is n. The exponential term in the above expressions is an approximation to $\exp\{-(\lambda X^* Y^* - \sum_{i=1}^{n} \lambda dx_i dy_i)\}$, i.e., represents the probability of no events within the region X^*Y^* but outside the incremental regions containing each event.

By conditioning on the occurrence of n events in the region which are distributed Poisson with parameter $\lambda X^* Y^*$,

$$f\{(x,y)_{(1)},\ldots(x,y)_{(n)}|n;\lambda\} = \frac{\lambda^n \exp(-\lambda X^*Y^*)}{\frac{(\lambda X^*Y^*)^n}{n!}\exp(-\lambda X^*Y^*)}, \quad n \geq 1$$

$$= \frac{n!}{(X^*Y^*)^n} \quad , \qquad (11)$$

which is the joint distribution for n bivariate uniform random variables ordered on one of the random variables as is shown in Appendix A. Note also the independence of the conditioned density from the parameter λ, i.e. the random variable N is a sufficient statistic for λ.

As a consequence of the above corollary, it is apparent that if the points of the HPPP, conditioned on the number of events observed to occur, are in fact ordered with respect to the increasing magnitude of the y-component, then no "information" is available about the ordering of the x-components, i.e., each of the n! orderings that can be induced on the x's by the orderings on the y's has probability 1/n!. This is readily apparent since in the bivariate uniform case the two components were independently selected. Hence, if $(x,y)_{(k)}$ is determined by $(x_k, y_{(k)})$, i.e. the points are labelled by the ordered y-component, then

$$\text{prob}\{X_k = X_{(j)}\} = \frac{1}{n} \qquad j = 1,2,\ldots,n$$

where $X_{(j)}$ is the j^{th} X_k in magnitude, and

$$\text{prob}\{X_1 = X_{(j)},\ j=1,\ldots,n;\ X_2 = X_{(k)},\ k=1,\ldots,j-1,j+1,\ldots n;$$
$$\ldots;\ X_n = X_{(\ell)}\} = \frac{1}{n!} \qquad (12)$$

Hence if the x-components of the points ordered on the y-components exhibit any natural ordering then the x- and y-components have not been independently selected and the observed process cannot be a HPPP. This will be the basis for many of the tests for a HPPP against a non-homogeneous planar Poisson process to be discussed later.

Lemma: If the bivariate process is Poisson and the regions are rectangular, then the projections of the events onto the coordinate axes may be shown to be univariate Poisson.

Proof: Consider the occurrence of events in a rectangular region of area X^*Y^*. Then by III the occurrence of an event in an incremental strip is independent of all occurrences outside the strip. Hence the projections onto the coordinate axes give rise to independent counts along the axes.

$$P_n(x,Y^*) = \frac{(\lambda Y^* x)^n \exp(-\lambda Y^* x)}{n!} \quad , \quad \begin{matrix} n = 0,1,\ldots \\ 0 \leq x \leq X^* \end{matrix}$$

and

$$P_n(X^*,y) = \frac{(\lambda X^* y)^n \exp(-\lambda X^* y)}{n!} \quad , \quad \begin{matrix} n = 0,1,\ldots \\ 0 \leq y \leq Y^* \end{matrix}$$

which gives the univariate Poisson distributions with parameters λY^* and λX^* respectively.

Note here the inherent dependence on the shape of the assumed regions. In using rectangular regions equal lengths in the marginals reflect equal areas in the bivariate distribution.

If the regions were circular then vertical projections onto the axes would represent decreasing area as the distance from the origin increased. Since the occurrence of events is assumed to be proportional to the area projected, an actual HPPP would induce a non-homogeneous process on the marginals due to the distortion in the mapping. For clarification, refer to Figure 2. However, if the regions are circular then radial projections could be utilized so that the event occurring at (x_0,y_0) in Figure 2 is represented in the x-marginal by an event at x_2. To define equal area projections in this case the transformation $x \rightarrow x^2 = x'$ is made, in which case a unit increase in x' defines the addition of a unit amount of area to the region. For example, if a unit area is generated by a circle of radius $r = 1$, then the area enclosed in the ring of $1 \leq r \leq \sqrt{2}$ is the unit area, as is the area in the ring $\sqrt{2} \leq r \leq \sqrt{3}$, etc. In general, $\sqrt{n} \leq r \leq \sqrt{n+1}$ defines in polar coordinates a ring with unit area.

Returning to the assumption of rectangular regions, three characteristics of the HPPP are now available which can be used as the basis for testing a sample for belonging to the HPPP description of events in a rectangular region R:

(A) Independence of the x-ordering from the ordering on the y-components.

(B) Univariate HPP (homogeneous Poisson process) in the x-marginal and, conditionally on n events in R, a

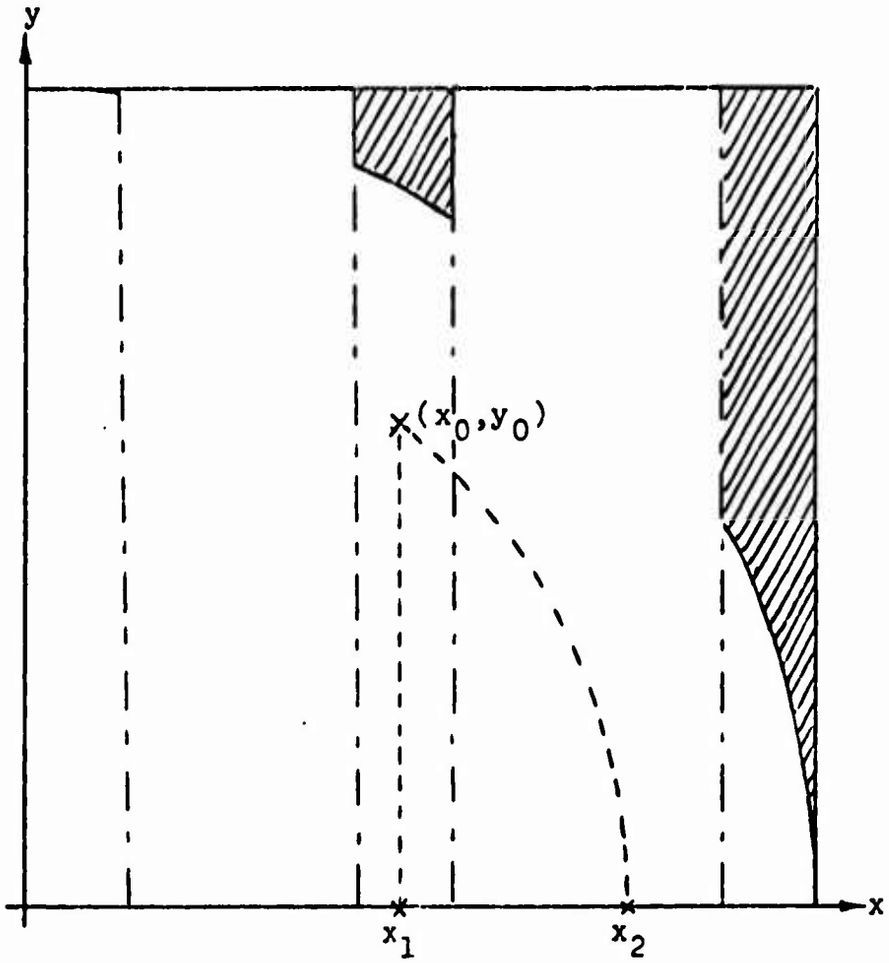

Figure 2. Vertical and radial projections of an event to form the marginal process. Shaded regions represent the deviations of projected areas arising from the rectangular projection of circular areas. Thus, the shaded regions indicate the degree of non-homogeneity induced by the mapping.

uniform distribution of the distances to events.

(C) Univariate HPP in the y-marginal and, conditionally on n events in R, a uniform distribution of the distances to events.

Property (A) can be tested against general alternatives using a rank correlation procedure (or Spearman's correlation, see Pearson and Hartley [1966, Table 44]). Properties (B) and (C) can be tested by standard univariate methods as in Cox and Lewis [1966].

Note that in the above discussion the interest lies in the nature of the process rather than in specifically describing the process. Thus the determination of the parameter λ of the Poisson process is not a current objective and it can be considered to be a "nuisance" parameter. Hence the conditioning argument above and the resulting independence of the tests from the value of the parameter are justifiable.

Now let α_A be the probability of a Type I error generated in testing for randomness, α_B be the corresponding probability in testing for HPPP in the x-marginal, and α_C likewise for the y-marginal. Then the probability of not falsely rejecting the HPPP hypothesis due to the randomness test is $1 - \alpha_A$, etc. Hence the combined probability of not falsely rejecting HPPP is 1 - prob {type I error} or

$$1 - P(I) = (1 - \alpha_A)(1 - \alpha_B)(1 - \alpha_C).$$

Therefore

$$P(I) = 1 - (1 - \alpha_A)(1 - \alpha_B)(1 - \alpha_C) \qquad (13)$$

is the probability of falsely rejecting a HPPP hypothesis.

If through physical considerations one of the tests seems more or less significant than the others, the analyst can choose the weightings to reflect the physical properties. Otherwise the values (and thus the tests) can be weighted equally. This need for the determination of weightings is the inherent disadvantage of a multi-level test.

The individual tests proposed above will be briefly described. For the rank correlation test, consider each x_i from $(x,y)_{(i)}$ which is ordered on the y-component. Also consider the ordered real values along the x axis, where $x_i = x_{(j)}$. Then

$$r_s = 1 - \frac{6 \sum_{i=1}^{n} (i - (j)_i)^2}{n(n^2 - 1)} , \qquad (14)$$

where $(j)_i$ is the position of x_i in the x-ordered sequence, is the rank correlation statistic.

The exact distribution for r_s can be approximated by fitting a distribution to its moments as discussed by Kendall and Stuart [1951, p.477]. The exact distribution of r_s is tabulated in Biometrika Tables For Statisticians [1966, Table 44, p.23] for observed values of n between 4 and 10, and the introduction to these tables gives approximations for $10 \leq n \leq 20$ and for $n > 20$. For $10 \leq n \leq 20$,

r_s can be treated as a product-moment correlation coefficient between normally distributed random variables. For $n > 20$, $r_s \sqrt{n-1}$ is assumed to be unit normal.

In testing the marginal distribution for HPP, two separate tests are proposed. First, the uniform conditional test is used to test against trends in the data. As stated in Cox and Lewis [1966, p. 153], "If the series has been observed for a fixed time t_o {length X*} and n events occur in $(o,t_o)\{(o,X^*)\}$, then the uniform conditional test is based on the variables $U_{(i)} = T_i/t_o$ $\{= X_{(i)}/X^*\}(i=1,\ldots n)$ conditionally on N_{t_o} being equal to n." The {brackets} are supplied to relate the material in Cox and Lewis [1966] to this specific problem, and $N_{t_o} = n$ means the number of occurrences observed is n. Note that in the conditioning of the realizations the "nuisance" parameters λX^* and λY^* are eliminated.

Secondly, a test based on the ordered inter-event spacings is used to test Poisson against stationary event processes which may be non-Poisson. For this test, Durbin's modifications of the uniform conditional test is used [Cox & Lewis, 1966, p. 155]. Referring to Figure 3, Durbin's modification describes a transformation from the random variable X to the random variable T and then to the random variable S.

Let $T_{n+1} = X^* - X_{(n)}$. If the $X_{(i)}$, $i = 1,2,\ldots,n$, describe the "times to events" in a Poisson process, then the T_i, $i = 1,2,\ldots,n+1$, are independent exponentially

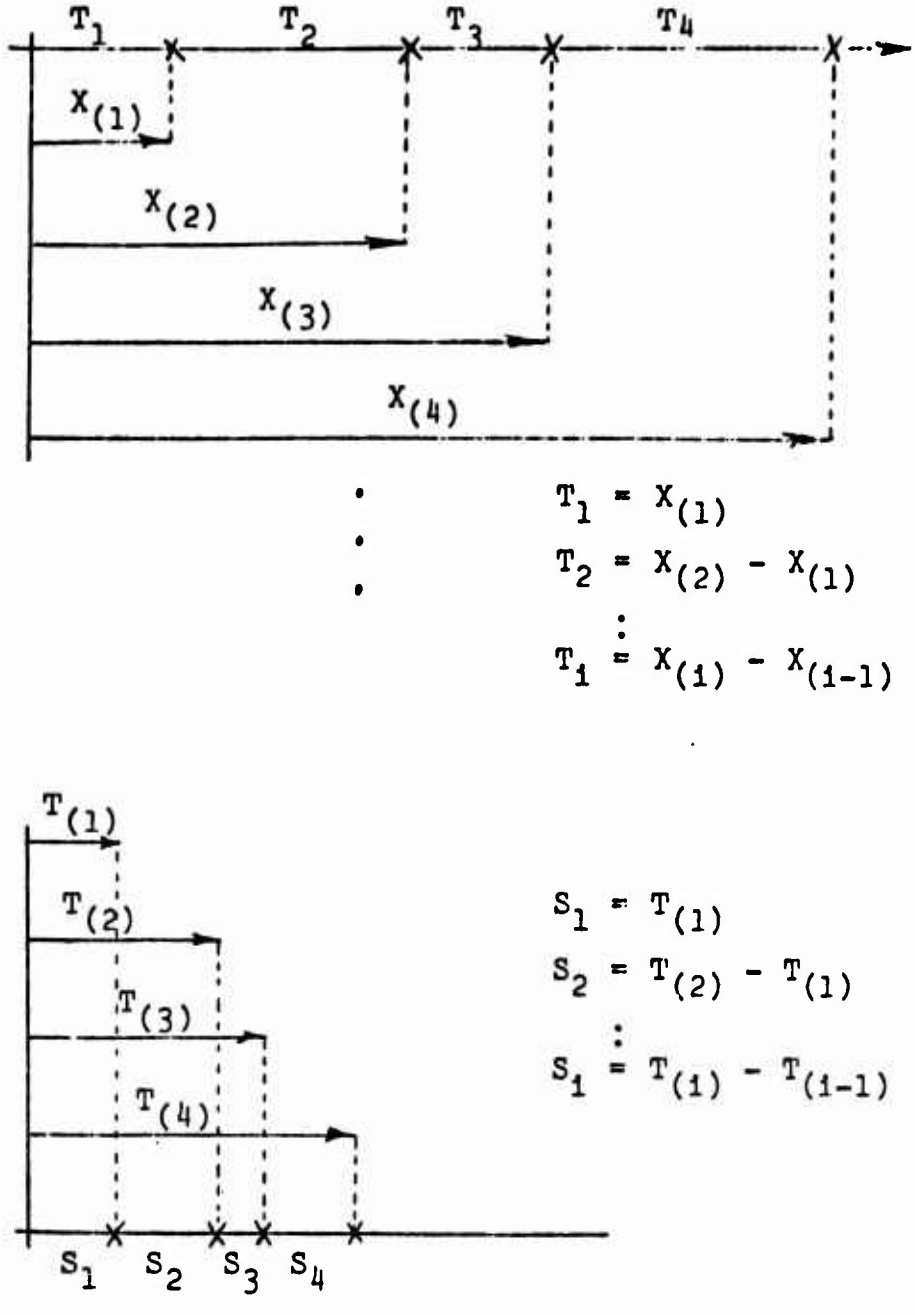

Figure 3. The generation of the transformed variables S_i from the original process $X_{(i)}$.

distributed random variables with parameter λ. If the T_i's are then ordered and the S_i's are generated as shown, then the S_i's are independent exponential random variables, where S_i has the expectation $1/((n+2-i)\lambda)$. Also the transformation $S_i' = (n+2-i)S_i$ defines independent identically distributed exponential random variables with parameter λ, and therefore $X_i' = \sum_{j=1}^{i} S_j'$, $i = 1,2,\ldots,n$ defines the times to events in a Poisson process with parameter λ, and $U_i' = \frac{1}{X_n'} \sum_{j=1}^{n} S_j'$ is the statistic upon which a new uniform conditional test is based.

Both tests should be utilized as the uniform conditional test is more powerful when testing for trends while Durbin's modification is relatively more powerful in testing against stationary event process alternatives. However, these tests are not independent of each other and thus cannot be combined as in (13).

As an alternative to the above procedure, the region of concern may be partitioned into several sub-regions and the number of events in each subregion used as a basis for χ^2 testing. This method is discussed by Kendall and Stuart [1951, pp. 574-5] who mention the problem of choosing the "right" partition, adding "Whether a particular partition has statistical interest depends on the purpose of the analysis". Due to the underlying uniformity of the conditional distribution, this problem reduces to the selection of the number of regions which are then used to form equal area sub-regions.

Another alternative to the above testing procedure is the evaluation of the sample product-moment correlation coefficient under the bivariate uniform distribution. The procedure is discussed by Kowalski [1972], but unfortunately his discussion does not address the bivariate uniform distribution. Kowalski makes two points very strongly: "Firstly, the distribution of r (the sample product-moment correlation coefficient under non-normal assumptions) may differ from its normal-theory form and, secondly, we may be in a situation in which ρ is a poor measure of association." Hence, if the exact distribution for r under the bivariate uniform distribution were known, then an exact test for the HPPP (given the occurrence of n events) could be devised.

Durbin [1970] has also proposed distance methods for testing bivariate distributions. The process herein described is well-suited to the methods Durbin uses since he first transforms the observations so that they occur uniformly on the unit square. Hence the natural transformation $x' = x/X^*$ and $y' = y/Y^*$ avoids the problem of possible lack of uniqueness which is the central objection to the use of distance methods. These methods allow the analyst to adopt Durbin's bivariate analog of the Kolmogorov-Smirnov tests. The advantage of this method is the elimination of difficulties concerning multi-level tests and partitioning tests.

The tests described in this section are very general in nature, i.e.

H_o: the process is HPPP is tested against

H_1: the process is not HPPP.

Hence the alternatives being tested against are multitudinous. If it is desired to test a realization as being from a HPPP against a specific form of departure from HPPP, better tests may be defined based on the nature of the specific alternative. For instance, one such departure could be non-homogeneity, i.e., where λ is not considered to be constant but rather a function of location; this subject is considered in chapter III. Another departure might be in the nature of the process itself. For example, events may occur according to some fixed plan in which case the process is deterministic and thus non-Poisson. A process may develop in which the occurrence of an event prohibits the occurrence of another event for some interval about itself, in which case events are not independent of other events and are thus non-Poisson.

It must be remembered, however, that tests against specific alternatives may ignore some features that a more general test would detect and thus each individual specific test applies only to the specific form of departure being considered.

Moreover, in all reasonable stationary alternatives, it does not seem possible to derive the likelihood function

of the observations. One thus cannot derive exact tests. For tests against specific alternatives based on distance methods, see Holgate [1972]. Tests based on spectra are discussed by Bartlett [1964].

C. SIMULATING A HPPP

Suppose one were concerned with searching for submarines which are assumed to be dispersed in such a manner that the locations at any moment are generated by a HPPP. If one search procedure is to be selected from many proposed search procedures, then a possible manner of comparing the effectiveness of the proposed procedures is to utilize each procedure against several simulated dispersions. In such a simulation, the only "variable" which would be of interest would be the procedures, so all variables such as detection and classification parameters, facilities available, etc., would remain constant. Another problem which might be considered would be the effect of the change of such parameters on the search procedure selected (i.e., a sensitivity analysis of the procedure to assumed operating characteristics and facilities).

By the initial remarks of Section B above and the statement of equation (12), several methods of artificially generating realizations of a HPPP can be determined. These methods may then be utilized to simulate the HPPP.

Assume that the parameter $\lambda X^* Y^*$ is given. To select the number N of events to be observed in the region with

area X*Y*, generate a random number U distributed uniformly on [0,1]. Set N = n if

$$\sum_{i}^{n-1} \frac{(\lambda X^*Y^*)^i \exp(-\lambda X^*Y^*)}{i!} < u \leq \sum_{i=1}^{n} \frac{(\lambda X^*Y^*)^i \exp(-\lambda X^*Y^*)}{i!}. \tag{15}$$

The summations can be evaluated using either χ^2 or Gamma Integral Tables [Cox and Lewis, 1966, p.24]:

$$\sum_{i}^{n-1} \frac{(\mu)^i e^{-\mu}}{i!} = \text{prob}\{\chi^2_{2n} > 2\mu\}$$

$$= \int_{\mu}^{\infty} \frac{v^{n-1} e^{-v}}{(n-1)!}\, dv.$$

Next, consider a random variable X distributed uniformly over (0,X*), denoted $X \sim U(0,X^*)$, and another independent random variable $Y \sim U(0,Y^*)$. As realizations of each random variable are generated, number them chronologically, i.e. in order of appearance. Generating n (as determined above) such realizations of each random variable yields 2 numbers: $x_1,\ldots,x_n,y_1,\ldots,y_n$.

The final problem remaining is to select a scheme for mating the x- and y- realizations to form ordered pairs which will constitute the realization of the HPPP. A few such schemes are enumerated:

1. The sequence $\langle(x_i,y_i)\rangle_{i=1}^{n}$ forms a HPPP.
2. If the y_i are ordered to form $\langle y_{(j)}\rangle_{j=1}^{n}$, then the sequence $\langle(x_i,y_{(i)})\rangle$ forms a HPPP.

3. Similarly, $\langle(x_{(i)},y_i)\rangle$ forms a HPPP.

4. Additionally, any random permutation of the x_i in 2, the y_i in 3 or either random variable in 1 can be used to form a HPPP. Thus $\langle(x_{n+1-i},y_{(i)})\rangle$ forms a HPPP, etc.

The goal of the simulation and the purpose of simulating the process as a part of the overall analysis must now be considered. If during the simulation it is desired to generate independent realizations of the process, then each iteration must involve a selection of n, the drawing of 2n uniform variates and the mating of the variates through some scheme such as those outlined in steps 1-4 above. On the other hand, if it is desired to utilize variance reduction techniques, then for any drawing of 2n random variates several schemes could be used for the mating process. Here independence is lost immediately and this loss must be balanced by some gain elsewhere in the analysis.

D. ESTIMATION AND TESTING FOR THE PARAMETER FROM A HOMOGENEOUS PLANAR POISSON PROCESS (HPPP)

If the hypothesis that the process is HPPP with some unknown value of the parameter λ is accepted, one might like to obtain a point estimate or confidence interval estimate for λ, or to test that the process has some given parameter λ_0. Note that the parameter λ, which was considered to be a nuisance parameter in the previous section where the structural aspects of the process <u>per se</u> were tested, now specifies the process completely.

Since, as was seen in Section B, it is possible to set up the joint probability density function of the observations in a HPPP, point estimation of λ can be based on the method of maximum likelihood. Note, however, that each observation consists of a single "look" at (or realization of) the process rather than n observations of a single random variable. Since it is a stochastic process the observations are not independent and identically distributed. Hence the usual justifications for maximum likelihood procedures are not valid; see Brown [1972] for extensions of maximum likelihood theory of estimation to realizations of a Poisson process.

Using the results of Brown [1972], suppose that n HPPP events are observed to occur in a rectangular region of area X^*Y^*. From (11), for $n \geq 0$,

$$L = f\{(x,y)_{(1)},\ldots,(x,y)_{(n)},n;\lambda\} = \lambda^n e^{-\lambda X^*Y^*} \qquad (16)$$

$$\ln L = n\ln\lambda - \lambda X^*Y^*, \quad (0 \leq \lambda < \infty.)$$

If $n \neq 0$, this function is $-\infty$ at $\lambda = 0$ and $\lambda = \infty$ and since $\frac{d \ln L}{d\lambda} = \frac{n}{\lambda} - X^*Y^*$, the slope of the function decreases monotonically from ∞ to $-X^*Y^*$. Thus ln L has a unique maximum at the point where $\frac{d \ln L}{d\lambda} = 0$. Setting this derivative equal to zero yields a unique maximum likelihood point estimate for λ as

$$\hat{\lambda} = \frac{n}{X^*Y^*}, \quad (n \geq 1) \qquad (17)$$

where $\hat{\lambda}$ is unbiased (since $E(\hat{\lambda}) = \frac{E(n)}{X^*Y^*} = \lambda$) and has variance λ/X^*Y^*. Note that as the observed area X^*Y^* becomes large, the variance of the estimate becomes small; thus, by Chebyshev's Inequality [Lamperti, 1966, p. 20]

$$P\{|\hat{\lambda} - \lambda| \geq a\} \leq \frac{\text{Var } \hat{\lambda}}{a^2}, \quad (a > 0)$$

and as $X^*Y^* \to \infty$,

$$P\{|\hat{\lambda} - \lambda| \geq a\} \to 0$$

for all positive a and hence $\hat{\lambda}$ converges to λ in probability. The latter statement is equivalent to the assertion that $\hat{\lambda}$ is a consistent estimator for λ. Also since the variance of $\hat{\lambda}$ is λ/X^*Y^*, $\hat{\lambda}$ has an estimated variance $\hat{\lambda}/X^*Y^* = n/(X^*Y^*)^2$, and an estimated standard error of $\sqrt{n}/X^*Y^*$.

If $n = 0$, the above method is not applicable. In thi case, it might be preferable to give a confidence interva estimate for λ. Specifically, a one-sided test alternative is used to generate a test for the assumed value λ_{null} using as an acceptance region only $n = 0$. Intuitively, λ_{null} will be small enough so that $\lambda_{null}X^*Y^* < 1$ (i.e., the expected number of observed events is less than 1). The hypothesis to be tested is H_0: $\lambda = \lambda_{null}$ vs. H_1: $\lambda > \lambda_{null}$. Defining a level of significance α from (16) by

$$\text{prob}\{N = 0|\lambda = \lambda_{null}\} = 1 - \alpha = e^{-\lambda_{null}X^*Y^*},$$

the hypothesis H_0 is accepted at the level α. Conversely, for any given value of α, λ_{null} may be determined by

$$- \lambda_{null} X^* Y^* = \ln (1 - \alpha)$$

$$\lambda_{null} = - \frac{\ln (1 - \alpha)}{X^* Y^*} ,$$

where the λ_{null} thus determined is the largest value of λ that the test will accept at the α level, given that $n = 0$.

Returning to the case of $n \geq 1$, in order to test that the parameter of the process has some given value λ_0, assume that n events from a HPPP are observed in a region of area X^*Y^*. The hypothesis to be tested is H_0: $\lambda = \lambda_0$ against the two-sided alternative H_1: $\lambda \neq \lambda_0$ although one-sided alternatives can also be considered. Since N is a random variable taking on all nonnegative integer values with some positive probability for any λ_0, there is always some possibility of an observed value of the random variable N (the observation being denoted n) falling outside any finite range of values. Thus a region (n^-, n^+) must be specified such that if N lies in the region the hypothesis H_0 is accepted; otherwise the hypothesis is rejected. The level α of the test is the probability, given $\lambda = \lambda_0$, that N falls outside the region (n^-, n^+).

Since the test has been defined to be two-sided, the level is split into upper and lower levels α^+ and α^-

so that $\alpha = \alpha^+ + \alpha^-$. The procedure must consider values of $\lambda < \lambda_0$ as well as values of $\lambda > \lambda_0$. To proceed, it is necessary to define

$$P_+(n^+;\lambda_0) = P\{N \geq n^+ | \lambda = \lambda_0\} = \alpha^+ \tag{18}$$

$$= \sum_{j=n^+}^{\infty} \frac{(\lambda_0 X^* Y^*)^j e^{-\lambda_0 X^* Y^*}}{j!} ,$$

and

$$P_-(n^-;\lambda_0) = P\{N \leq n^- | \lambda = \lambda_0\} = \alpha^- \tag{19}$$

$$= \sum_{j=0}^{n^-} \frac{(\lambda_0 X^* Y^*)^j \exp(-\lambda_0 X^* Y^*)}{j!} .$$

Thus, for a given α^+, an n^+ may be determined such that the statement (18) just holds. Also, for a given α^-, a n^- may be determined such that (19) just holds.

The null hypothesis is accepted at the α level if the observed value of N falls between the two prescribed limits ($n^+ > n^-$), where prob$\{N \notin (n^-,n^+)\} = \alpha$. Note that as stated, the result is indeterminate since α, once given, leads to many values for α^+ and $\alpha^- = \alpha - \alpha^+$ which satisfy the given α. The manner of selecting α^+ and α^- must be stated. Arbitrarily it may be desirable to set $\alpha^+ = \alpha^- = \alpha/2$. Asymptotically, this choice of a symmetric acceptance region is reasonable since as n increases, the distribution of N

is approaching the (symmetric) normal distribution. The choice of equal α^+ and α^- may not be reasonable, however, for small $\lambda_0 X^* Y^*$ since the Poisson distribution is positively skewed.

The statement $\text{prob}\{N \notin (n^-, n^+) | \lambda = \lambda_0\} = \alpha$ is the result of the test of the hypothesis H_0: $\lambda = \lambda_0$ at a given, fixed level α. It is this result from which one must usually draw conclusions regarding specification of the process.

If the information thus available, i.e. H_0 is rejected or accepted at the pre-determined α level, is deemed insufficient for the purposes of a decision maker (for example) then another possibility is that the post-analysis information may be extended by determining for each observation the exact α, α_e, at which the hypothesis would have been rejected. The decision maker is then left with the problem of the determination of his own level of significance, possibly based on his intuitive grasp of the problem and its significance in a larger frame of reference. Once he has determined his preferred significance level, the hypothesis is rejected or accepted at the specified level by comparison with α_e. Thus the decision maker has gained some influence over the analysis but has had to pay with some time to reflect on the problem at hand. Alternatively, he can use α_e <u>informally</u> as a "goodness of fit" of the hypothesis.

Using (18) and (19), the significance test is defined conventionally [Cox and Lewis, 1966, p. 30] to be: the

hypothesis $\lambda = \lambda_0$ would be accepted at the level of significance α in a two-sided equi-tailed test if the observed number of events, n, is such that n, when used alternatively in (18) and (19) (i.e., is assumed to be one or the other of the end-points of the acceptance region), produces α_e as a solution to

$$P(n;\lambda_0) = 2\min(P_+(n;\lambda_0), P_-(n;\lambda_0)) = \alpha_e. \qquad (20)$$

Note that each observed value of n generates a new α_e for any assumed λ_0, hence $\alpha_e = \alpha_e(n, \lambda_0)$. For example, $P(30;20) = .0436$, $P(20;20) = .7628$ and $P(15;20) = .1352$.

It can be seen that the fixed level procedure is computationally simpler, since for a specified α and λ_0, the interval (n^-, n^+) need only be computed once while in the latter procedure α_e must be recomputed following each observation of N.

The inverse of the above approach which utilized the two-sided equi-tailed test of significance for a given value λ_0 leads to the determination of confidence interval estimates of λ. Given that n events are observed, it is required to determine some limits on the range of λ such that the true parameter value λ^* lies within the stated limits with a probability $1 - \alpha$. That is, it is required to establish a $\lambda^-(N)$ and a $\lambda^+(N)$ such that

$$P(\lambda^-(N) \leq \lambda^* \leq \lambda^+(N) | N = n) = 1 - \alpha. \qquad (21)$$

Using $P\{N \leq n | \lambda = \lambda^+\} \leq 1 - \alpha^+$ to define a λ^+ as the greatest λ such that equality just holds and similarly using $P\{N \leq n | \lambda = \lambda^-\} = \alpha^-$ to define a λ^- establishes the limits such that (21) holds. For a proof of this, see Brownlee [1965, p. 121]. Note that for each realization of N, a new ordered pair (λ^-, λ^+) is defined so that the ordered pair is a function of a random variable and hence is itself a random interval. The procedure only states that for $(1 - \alpha) \times 100\%$ of the observations the true parameter λ^* will lie within the limits selected. The limits for observed n from 0 to 50 are tabulated [Pearson and Hartley, 1966, Table 40].

For a normal approximation to the confidence interval, Cox and Lewis [1966, p. 31] define the upper α point of the unit normal distribution as c_α, and give the relationship

$$\text{prob}\{-c_{\frac{1}{2}\alpha} < \frac{N - \lambda X^* Y^*}{(\lambda X^* Y^*)^{1/2}} \leq c_{\frac{1}{2}\alpha}\} = 1 - \alpha,$$

the relationship being correct as $\lambda X^* Y^* \to \infty$. The confidence limits thus obtained are, to a second degree of approximation using a continuity correction and the estimate $\sigma(\hat{\lambda}) = \sqrt{n}/X^* Y^*$,

$$n + \frac{1}{2}c^2_{\frac{1}{2}\alpha} \pm c_{\frac{1}{2}\alpha}\sqrt{n}.$$

For example, if 50 events are observed from a HPPP, the exact .05 confidence interval is $37.11 \leq \lambda X^* Y^* \leq 65.92$

[Pearson and Hartley, 1966, Table 40] whereas the normal approximation gives $37.79 \leq \lambda X^*Y^* \leq 66.07$.

There also exist χ^2 approximations to the significance tests and confidence intervals [Cox and Lewis, 1966, p. 33; Brownlee, 1965, p. 173].

III. NON-HOMOGENEOUS PLANAR POISSON PROCESSES (NHPPP)

A. GENERAL DISCUSSION

If the stochastic process described above is generalized to allow the probabilistic structure of the event process to be dependent on the location of the events, a non-homogeneous planar process is evidenced. In the simplest such case a non-homogeneous planar Poisson process (NHPPP) arises if, in the definition of the Poisson process given above, assumption I is modified to become

I". There exists a positive finite function $\lambda(x,y) > 0$.

Also note that II is changed by the fact that the number of events in any region is not only a function of the area of the region, but also depends on the location of that region within the universe under consideration. Thus λ is now expressed as $\lambda = \lambda(x,y)$, and assumption II becomes

II". $\text{prob}\{N(R_i) = n\}$

$$= \frac{\{\Lambda(A_i)\}^n \exp\{-\Lambda(A_i)\}}{n!}$$

where $\Lambda(A_i) = \int_{A_i} \lambda(x,y)\, dxdy$ the symbol $\int_{A_i}$ implying the integral over an area and $\lambda(x,y)$ is assumed to be continuous over R_i (with area A_i) so that the integral is valid.

Assumption III remains unmodified, i.e. events occur independently of any other event or collection of events.

Under the additional assumption that $\lambda(x,y)$ is continuous within the region of consideration, the incremental

development of Chapter II may be extended to achieve a description of the NHPPP. Additionally the continuity assumption on λ and the definition of the parameter in the process as an integral over λ eliminates the difficulties of line discontinuities, although there may be cases where this is an important component of the problem. This problem is not considered here.

Referring back to Figure 1 in Section II-A, consider specifically the incremental strip defining region R_1. If the strip is divided into n sub-regions of equal area by taking n equal increments along the x direction each of length δx, then, under the assumptions on the behavior or $\lambda(x,y)$, the process in the ith sub-region can be approximated by a HPPP with parameter $\lambda = \lambda(x,y)_i$ where $(x,y)_i$ is an arbitrary point in the ith sub-region. Specifically (and arbitrarily) the lower left point is chosen for the succeeding discussion; thus the parameter for the first sub-region has parameter $\lambda = \lambda(0,Y^*)$. Continuing, the probability statements for occurrence of events become

$$P_1(x,y) = \lambda(0,Y^*)\delta x\Delta y + o(\delta x,\Delta y) \quad 0 \leq x \leq \delta x,$$

$$P_1(x,y) = \lambda(\delta x,Y^*)\delta x\Delta y + o(\delta x,\Delta y) \quad \delta x \leq x \leq 2\delta x,$$

$$\vdots$$

$$P_1(x,y) = \lambda(j\delta x,Y^*)\delta x\Delta y + o(\delta x,\Delta y) \quad j\delta x \leq x \leq (j+1)\delta x,$$

where $j = 0,1,\ldots,n-1$, $Y^* \leq y \leq Y^*+\Delta y$ and $n\delta x = X^*$.

Since the probability of more than one event in R_1 is $o(X^*\Delta y)$ the probability statements above are additive and

$$\text{prob \{one event in } R_1\} = \sum_{j=0}^{n-1} \lambda(j\delta x, Y^*)\delta x \Delta y + o(X^*\Delta y).$$

In the limit as $n \to \infty$, by the definition of an integral

$$\text{prob \{one event in } R_1\} = \{\int_0^{X^*} \lambda(x,Y^*)dx\}\Delta y + o(X^*\Delta y).$$

By similar argument,

$$\text{prob \{one event in } R_2\} = \{\int_0^{Y^*} \lambda(X^*,y)dy\}\ \Delta x + o(Y^*\Delta x)$$

and

$$\text{prob \{one event in } R_3\} = \lambda(X^*,Y^*)\Delta x\Delta y + o(\Delta x\Delta y).$$

By comparison with equations (3), (4) and (5) the above statements lead to definitions for average parameters for each of the regions R_1, R_2 and R_3 as

$$\bar{\lambda}_1(X^*,Y^*) = \frac{1}{X^*}\int_0^{X^*} \lambda(x,Y^*)dx,$$

$$\bar{\lambda}_2(X^*,Y^*) = \frac{1}{Y^*}\int_0^{Y^*} \lambda(X^*,y)dy, \qquad (22)$$

and

$$\bar{\lambda}_3(X^*,Y^*) = \lambda(X^*,Y^*).$$

Using these average parameters, the equations (3), (4) and (5) are generalized, resulting in the following statements:

$$\text{prob \{one event in } R_1\} = \bar{\lambda}_1(X^*,Y^*)X^*\Delta y + o(X^*\Delta y),$$

$$\text{prob \{one event in } R_2\} = \bar{\lambda}_2(X^*,Y^*)Y^*\Delta x + o(Y^*\Delta x), \quad (23)$$

and

$$\text{prob \{one event in } R_3\} = \bar{\lambda}_3(X^*,Y^*)\Delta x\Delta y + o(\Delta x\Delta y).$$

Using the result (22) as defining the parameter in each of the incremental areas in Figure 1, equations (6), (7) and (8") become

$$P_n(X^*+\Delta x,Y^*) = P_n(X^*,Y^*)[1-\bar{\lambda}_2Y^*\Delta x]+P_{n-1}(X^*,Y^*)\bar{\lambda}_2Y^*\Delta x + o(Y^*\Delta x), \quad (24)$$

$$P_n(X^*,Y^*+\Delta y) = P_n(X^*,Y^*)[1-\bar{\lambda}_1X^*\Delta y]+P_{n-1}(X^*,Y^*)\bar{\lambda}_1X^*\Delta y + o(X^*\Delta y), \quad (25)$$

and

$$\begin{aligned} P_n(X^*+\Delta x,Y^*+\Delta y) = {} & P_n(X^*,Y^*+\Delta y) + P_n(X^*+\Delta x,Y^*) - P_n(X^*,Y^*) \\ & - \bar{\lambda}_3\Delta x\Delta y[P_n(X^*,Y^*) - P_{n-1}(X^*,Y^*] \quad (26) \\ & + \bar{\lambda}_1\bar{\lambda}_2X^*Y^*\Delta x\Delta y[P_n(X^*,Y^*)-2P_{n-1}(X^*,Y^*) \\ & + P_{n-2}(X^*,Y^*)] \\ & + o(Y^*\Delta x) + o(X^*\Delta y) + o(\Delta x\Delta y). \end{aligned}$$

Rearranging terms, dividing by $\Delta x \Delta y$ and taking the double limit as $\Delta x \to 0$ and $\Delta y \to 0$ yields

$$\frac{\partial^2 P_n(X^*,Y^*)}{\partial x \partial y} = -\bar{\lambda}_3[P_n(X^*,Y^*) - P_{n-1}(X^*,Y^*)] \qquad (27)$$

$$+ \bar{\lambda}_1\bar{\lambda}_2 X^*Y^*[P_n(X^*,Y^*) - 2P_{n-1}(X^*,Y^*) + P_{n-2}(X^*,Y^*)]$$

which, together with the boundary condition that P_n is a probability statement, gives

$$P_n(X^*,Y^*) = \frac{(\Lambda(X^*,Y^*))^n}{n!} \exp\{-\Lambda(X^*,Y^*)\}, \quad n = 0,1,\ldots \qquad (28)$$

where

$$\Lambda(X^*,Y^*) = \int_0^{X^*}\int_0^{Y^*} \lambda(u,v)\, dudv. \qquad (29)$$

Thus, the number of events occurring in a region bounded by the coordinate axes, $x = X^*$ and $y = Y^*$ has a Poisson distribution with mean given by $\Lambda(X^*,Y^*)$. The mean can be considered to reflect the cumulative effect of $\lambda(x,y)$ in the region of concern.

If n events from a NHPPP are observed to occur in a rectangular region defined as usual with area X^*Y^*, and the events occur at $(x,y)_{(i)}$, $i = 1,\ldots,n$, the labelling done on the magnitude of the y-component, then the joint density of the events and the probability that the number of events in X^*Y^* is n is given by

$$\prod_{j=1}^{n} \lambda((x,y)_{(j)})\exp\{-\Lambda(X^*,Y^*)\}. \qquad (30)$$

Note that (30) is a direct generalization of (16).

Hence the NHPPP can be described in a fashion similar to the HPPP, but the expressions have acquired increased complexity due to the necessity for the inclusion of integrals to define the parameters. The degree of added complexity is dependent upon the choice of the specific functional form for $\lambda(x,y)$. The next section develops the expressions for one specific form.

B. A SPECIAL CASE

To consider the location dependent type of process, a particular form for $\lambda(x,y)$ is chosen as

$$\lambda(x,y) = \exp\{\alpha + \beta x + \gamma y + \delta xy\}. \qquad (31)$$

Note that if $\beta x + \gamma y + \delta xy$ changes very little over the range of interest of x and y, then

$$\lambda(x,y) \tilde{=} (1 + \beta x + \gamma y + \delta xy)\exp\{\alpha\}. \qquad (32)$$

Other relationships may be used; however, they may cause necessary and untidy restrictions on the values which the constants α, β, γ and δ may assume. In particular, $\lambda(x,y)$ must be greater than 0 and the bivariate exponential

polynomial (31) ensures this with no restrictions on the range of the parameters.

Additionally, algebraic manipulation of the form reveals that the curves of $\lambda(x,y) = c$, c a constant, include some interesting properties.

1. If $\delta = 0$, then $\ln \lambda(x,y) = c$ is a family of straight lines in the plane, intersecting the x-axis at an angle $\theta = \tan^{-1}(-\beta/\gamma)$. In this case a clock-wise rotation of the coordinate axes through an angle θ would give an exponential function of y only.

2. If $\delta \neq 0$, then $\ln \lambda(x,y) = c$ describes a system of contour lines which form a hyperbolic paraboloid with a saddlepoint at $(-\gamma/\delta,-\beta/\delta)$ as is illustrated in Figure 4. It may be helpful to interpret the Figure in terms of a section of forest which has been sampled. The line r describes a possible direction of steepest ascent (DSA) which passes through or near the region being sampled. This DSA may not be a topographic feature but rather a mathematical expression for a possible increase in density of trees along some line. Obviously, there may exist a strong correlation between this mathematical DSA and some topographic features. Note here that along the DSA maximal values for $\lambda(x,y)$ are found in the sense that departing the DSA at right angles leads to decreased values for $\lambda(x,y)$, i.e. decreases in the forest density.

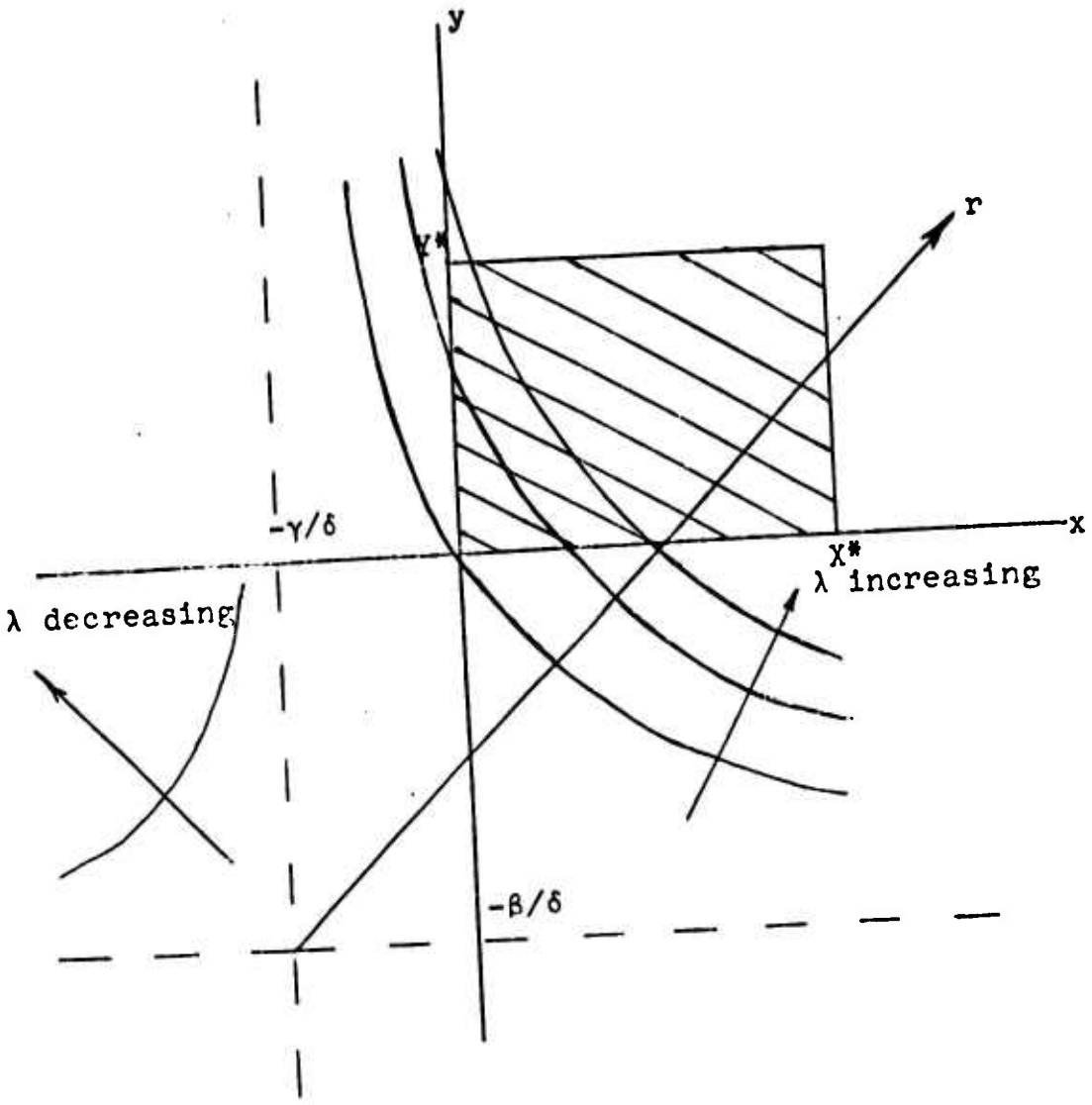

Figure 4. Contour lines for $\lambda(x,y) = \exp\{\alpha+\beta x+\gamma y+\delta xy\}$. Note asymptotes and region being described (hatched.) Here $\beta/\gamma = 2$, $\beta/\delta = 4$ and all coefficients are positive.

3. The exponential form can be extended with little conceptual difficulty, but possibly greatly increased mathematical difficulty, to describe a much wider range of possible circumstances. For instance, it is reasonable to assume that the DSA line will bend; hence terms such as $\varepsilon x^2 y$ and $\xi x y^2$ and higher order may be included in the exponent.

For the special form of (31), the cumulative or integrated intensity function $\Lambda(X^*,Y^*)$ is given by equation (29) and becomes

$$\Lambda'(X^*,Y^*) = \frac{\Lambda(X^*,Y^*)\cdot\delta}{\exp(\alpha-\beta\gamma/\delta)} = \overline{Ei}\{(\gamma+\delta X^*)(\tfrac{\beta}{\delta}+Y^*)\}-\overline{Ei}\{(\gamma+\delta X^*)\tfrac{\beta}{\delta}\} \tag{33}$$

$$- \overline{Ei}\{(\beta+\delta Y^*)\tfrac{\gamma}{\delta}\} + \overline{Ei}\{\tfrac{\beta\gamma}{\delta}\},$$

where $\overline{Ei}(\cdot)$ is the exponential integral and $\overline{Ei}(\cdot) = c + \ln(\cdot) + \sum_{i=1}^{\infty} \frac{(\cdot)^i}{i!\, i}$, where $c = .577216$ is a constant, as defined in Jahnke and Emde [1945, p. 2].

The likelihood function for the NHPPP may be developed in a manner similar to that used in the discussion of the HPPP. The discussion leading up to (16) is modified by the fact that the parameter is location dependent resulting in

$$L = \exp\{-\Lambda(X^*,Y^*)\} \prod_{i=1}^{n} \lambda((x,y)_{(i)}), \quad (n \geq 1) \tag{34}$$

where $(x,y)_{(i)}$ is a labelling of the coordinates of the n point events. Thus,

$$\ln L = -\Lambda(X^*Y^*) + \sum_{i=1}^{n} \ln \lambda((x,y)_{(i)})$$

follows directly. For the special case of $\lambda(x,y)$ given by (31), equation (34) becomes

$$\ln L = -\Lambda(X^*Y^*)+\beta \sum_{i=1}^{n} x_i+\gamma \sum_{i=1}^{n} y_i+\delta \sum_{i=1}^{n} x_i y_i+n\alpha \qquad (35)$$

where $\Lambda(X^*Y^*)$ is given by (33).

The above joint density, or likelihood, function provides a functional form which may be manipulated to accomplish the two principal concerns of the analysis of point processes: hypothesis testing and parameter estimation. The obvious null hypothesis is H_0: $\beta = \gamma = \delta = 0$, in which case the nonhomogeneous Poisson process is being tested for homogeneity since a non-zero α yields a constant parameter $\lambda = \exp(\alpha) > 0$. Should the above null hypothesis be rejected, then the analysis proceeds to develop estimates for the parameters β, γ and δ. This phase of the analysis may proceed differently depending on how many and which of the parameters were tested as being different from zero. The complete, and most complicated, situation develops when all parameters are determined to be non-zero. Testing of parameters is the topic of Chapter IV while Chapter V discusses the estimation of parameters determined to be non-zero as a result of the testing procedure.

IV. TESTING FOR NON-ZERO PARAMETERS

A. PRELIMINARIES

It is desired to formulate a method for testing the data (i.e., the number of events and their locations) in order to determine which of the parameters in the model given by (35), specifically α, β, γ and δ, are non-zero. Note that three assumptions are inherent at the outset: first, that the NHPPP model is valid; second, that the testing for homogeneity in Section II-B led to the rejection of the hypothesis of homogeneity; and third, that the physical phenomena can be modelled by the NHPPP given by (34) with the parameter $\lambda(x,y)$ given by (31).

Testing the Poisson hypothesis per se when the function $\lambda(x,y)$ is not known is a compound problem which will not be considered here. It is analagous to the compound probler in regression analysis of testing both for an unknown regres ion function and for independent equal variance errors.

From the third assumption, the likelihood function for the data is given by

$$L = \exp\{-\Lambda(X^*,Y^*)\}\xi^n \exp\{\beta\Sigma x_i + \gamma\Sigma y_i + \delta\Sigma x_i y_i\} , \quad (36)$$

where, for clarity in the future development, $\xi = \exp\{\alpha\}$. Conditioning on the occurrence of n events, $n \geq 1$ and defining $L\{(x,y)_{(1)},\ldots,(x,y)_{(n)}|n;\lambda(x,y)\} = L(n)$ leads to

$$L(n) = \frac{L}{\text{prob}(N=n)} = \frac{n!\,\exp(\beta\Sigma x_i + \gamma\Sigma y_i + \delta\Sigma x_i y_i)}{\left(\int_0^Y \int_0^X \exp(\beta u + \gamma v + \delta uv)\,du\,dv\right)^n} \tag{37}$$

where L(n) is read "the likelihood function conditioned on the occurrence of n events." Note that conditioning on the number of events observed has resulted in an expression which is independent of the parameter ξ(or α), i.e. for given β, γ, and δ, n is a sufficient statistic for α. This is convenient because α here is a "nuisance" parameter since the terms of interest are those which would indicate non-homogeneity rather than the establishment of the overall rate of occurrence. Thus by using the conditional likelihood α may be eliminated and the testing can proceed for non-zero β, γ and δ. In other words the value of α should not influence the testing for non-homogeneity parameters.

If $n = 0$, certainly no departure from homogeneity could be evidenced and hence this case is covered by HPPP; see II-B above. Hence the case of interest is $n \geq 1$.

Physically, the model (35) gives rise to a parameter surface $\lambda(x,y)$ which has the properties:

(a) $\beta \neq 0$; $\gamma \neq 0$; $\delta \neq 0$: ln λ forms a hyperbolic paraboloid superimposed on a tilted plane, i.e. some "warping" of the tilted plane is evidenced.

(b) $\beta \neq 0$; $\gamma \neq 0$; $\delta = 0$: ln λ forms a plane, tilted with respect to the x-y plane.

(c) $\beta = \gamma = 0$; $\delta \neq 0$: ln λ forms a hyperbolic paraboloid.

(d) $\beta = \gamma = \delta = 0$: ln λ forms a plane parallel to the x-y plane, i.e. a HPPP is evidenced.

There are a number of possibilities for testing:

(a) A test of

H_0: $\beta = \gamma = \delta = 0$

against

H_1: at least one of the parameters β, γ, $\delta \neq 0$

is a test for non-homogeneity which is more specific than those in Section II-B and is easily derived by likelihood ratio techniques.

(b) The above test is not of great interest; generally the specific non-zero parameter is desired rather than just that at least one of the three is non-zero. This leads to the question of selecting the significant subset, a problem which is difficult and as yet is unresolved.

(c) The simpler problem is to assume an ordering, i.e. that if $\beta = \gamma = 0$, the process is homogeneous (δ is then assumed to be 0) and if β or γ is non-zero but $\delta = 0$, then higher order terms are assumed to be zero. However, if the test indicates non-zero β or γ this may be due to an aliasing effect because of a non-zero δ. If further testing of $\delta = 0$ against $\delta \neq 0$ reveals $\delta \neq 0$, then it may well be that the true situation is $\beta = \gamma = 0$ but $\delta \neq 0$. The procedure to be followed will not discriminate this case.

The same aliasing effect occurs in testing of $\delta = 0$ against $\delta \neq 0$ where β and γ are non-zero and it is desirab' to perform this test without the effects of the non-zer β and γ. These are thus nuisance parameters, as was the case with α in testing β and γ. For the present model (35), one can eliminate these parameters because it is seen from the exponential form (36) that for any δ, $(n, \Sigma x_i, \Sigma y_i)$ is a set of sufficient statistics for (α, β, γ). Thus $\delta = 0$ is tested with some function of $\Sigma x_i y_i$ given n, Σx_i and Σy_i. This statistic has a distribution independent of the parameters α, β, γ.

The reason for basing the conditional test on $\Sigma x_i y_i$ is that this is (conditionally) a sufficient statistic for δ.

B. SPECIFIC TESTS

Assuming that some ordering exists on the parameters s discussed in possibility (c) above, tests are performed using the sufficient statistics $(n, \Sigma x_i, \Sigma y_i, \Sigma x_i y_i)$ to determine if any non-homogeneity is evidenced by the data (i.e., through the statistics). This testing is more specific in nature than the testing encountered in Section II-B above due to the selection of a particular model. The set of sufficient statistics arises from this choice of a specific model to use as an alternative to homogeneity.

The testing will assume the following sequence:

(i) Condition on n and set $\delta = 0$. Test $H_{0(i)}: \beta = \gamma = 0$ against $H_{1(i)}: \beta \neq 0$ or $\gamma \neq 0$. Note that it would not be informative to test either β or γ as a separate entity since in the formulation of the model β and γ are unique only up to an angle of rotation. That is, testing of β and γ jointly amounts to the detection of any tilt in $\ln \lambda(x,y)$ with respect to the x-y plane, regardless of the direction of the tilt. Failure to reject $H_{0(i)}$ leads to the assumption of homogeneity due to the assumed ordering.

(ii) Rejection of $H_{0(i)}$ leads to testing of

$$H_{0(ii)}: \quad \delta = 0, \ -\infty < \beta < \infty \quad \text{and} \quad -\infty < \gamma < \infty$$

against

$$H_{1(ii)}: \quad \delta \neq 0; \ -\infty < \beta < \infty \quad \text{and} \quad -\infty < \gamma < \infty.$$

The test thus specifies γ and β as nuisance parameters. In this test it is necessary to first condition on n, Σx_i and Σy_i to eliminate the nuisance parameters.

In (i), conditioning on n and setting $\delta = 0$ leads to

$$L(n) = \frac{(\beta\gamma)^n \, n! \, \exp\{\beta \Sigma x_i + \gamma \Sigma y_i\}}{(\exp\{\beta X\} - 1)^n \, (\exp\{\gamma Y\} - 1)^n} .$$

From this it is seen that the statistics $(\Sigma x_i, \Sigma y_i)$ are (conditionally) jointly sufficient statistics for β and γ.

Under $H_{0(i)}$,

$$\Sigma x_i/n \rightarrow N(X^*/2, X^{*2}/12n) \quad \text{and} \quad \Sigma y_i/n \rightarrow N(Y^*/2, Y^{*2}/12n)$$

and the statistics are independent (see Section II-B). Hence the expression

$$\left(\frac{\Sigma x_i/n - X^*/2}{X^*/\sqrt{12n}}\right)^2 + \left(\frac{\Sigma y_i/n - Y^*/2}{Y^*/\sqrt{12n}}\right)^2$$

is asymptotically χ^2_2. Rejection or acceptance of $H_{0(i)}$ is based on the adherence of the calculated value of this sum to the χ^2 distribution , i.e. H_0 is accepted if this sum has sufficiently small values. Acceptance of $H_{0(i)}$, as stated earlier, leads to assumption of HPPP; refer to Chapter II.

Following the rejection of $H_{0(i)}$ it is necessary to proceed with testing of $H_{0(ii)}$. As can be seen from an examination of (37), the complexity of the exact distribution following another conditioning argument (i.e. on n, Σx_i and Σy_i) is prohibitive. However, for large sample sizes the conditional distribution can be approximated from the fact that $\Sigma x_i/n$, $\Sigma y_i/n$ and $\Sigma x_i y_i/n$, conditioned on n, are jointly normally distributed for large n. Thus the asymptotic distribution of $\Sigma x_i y_i/n$, given n, $\Sigma x_i/n$ and $\Sigma y_i/n$, can be found from normal theory multiple regression results.

Under the assumption that $\beta = \gamma = \delta = 0$, the trivariate normal distribution which arises is characterized by a vector and a matrix. The vector ($\underline{\mu}$) of expected values and the variance-covariance matrix ($\underline{\Sigma}$) are given by

$$\underline{\mu} = \begin{pmatrix} X/2 \\ Y/2 \\ XY/4 \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_{xy} \end{pmatrix}$$

and

$$\underline{\Sigma} = \begin{pmatrix} X^2/12n & 0 & X^2Y/24n \\ 0 & Y^2/12n & XY^2/24n \\ X^2Y/24n & XY^2/24n & 7X^2Y^2/144n \end{pmatrix}$$

from which $\rho_{12} = 0$ and $\rho_{13} = \rho_{23} = 0.65465$.

In the model given above,

$$H_{0(ii)}: \quad \delta = 0; \quad -\infty < \beta < \infty; \quad -\infty < \gamma < \infty$$

is to be tested against

$$H_{1(ii)}: \quad \delta \neq 0; \quad -\infty < \beta < \infty; \quad -\infty < \gamma < \infty.$$

Since $\Sigma x_i y_i$ is a sufficient statistic for δ when n, Σx_i and Σy_i are given, the test can be based on $\Sigma x_i y_i$. Its asymptotic (conditional) normal distribution has mean μ'_{xy} and standard deviation σ'_{xy} given by

$$\mu'_{xy} = \mu_{xy} - \frac{\sigma_{xy}}{\sigma_x} \rho_{13} (\Sigma x_i/n - X/2) +$$

$$+ \frac{\sigma_{xy}}{\sigma_y} \rho_{23} (\Sigma y_i/n - Y/2)$$

and

$$\sigma'_{xy} = \sigma_{xy} (1 - \rho_{13}^2 - \rho_{23}^2)^{\frac{1}{2}} .$$

Thus under $H_{0(ii)}$, $\frac{\Sigma x_i y_i/n - \mu'_{xy}}{\sigma'_{xy}}$ is distributed as a unit normal variate and $H_{0(ii)}$ is accepted if this statistic has sufficiently small values. Failure to reject $H_{0(ii)}$ would imply that the $\ln \lambda(x,y)$ plane is tilted with respect to the x-y plane, but no "warping" is evidenced.

The above development relies heavily on asymptotic assumptions. Small sample problems will be much more difficult to analyze. Any point in the above procedure which lead to rejection of any hypothesis would require the analysis to proceed with the estimation of the non-zero parameters. This is the subject of the next chapter.

V. ESTIMATION OF PARAMETERS

It is desired to formulate a method for estimating the parameters α, β, γ and δ of the non-homogeneous planar Poisson model given in IV-A where it has been established that a non-homogeneous process is evidenced by the data.

Taking the logarithm of the conditional likelihood function (37) results in

$$\ln L(n) = \ln n! + \beta\Sigma x_i + \gamma\Sigma y_i + \delta\Sigma x_i y_i + n \ln A, \quad (38)$$

where $A = \Lambda(X^*,Y^*)/\xi$. Point estimation of $(\alpha, \beta, \gamma, \delta)$ by the method of maximum likelihood uses the conditional likelihood function (38) to develop the estimates. See Section II-D for comments regarding use of maximum likelihood in this application. The solution to the set of simultaneous equations

$$\begin{aligned}
\Sigma x_i - \frac{n}{A}\int_0^{Y^*}\int_0^{X^*} u \exp\{\beta u + \gamma v + \delta uv\}\, dudv &= 0 \\
\Sigma y_i - \frac{n}{A}\int_0^{Y^*}\int_0^{X^*} v \exp\{\beta u + \gamma v + \delta uv\}\, dudv &= 0 \qquad (39) \\
\Sigma x_i y_i - \frac{n}{A}\int_0^{Y^*}\int_0^{X^*} uv \exp\{\beta u + \gamma v + \delta uv\}\, dudv &= 0,
\end{aligned}$$

if obtainable, provides the point estimates $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\delta}$. Note that this approach neglects the homogeneous term

during the estimation of the parameters giving rise to non-homogeneity. The neglected parameter may be estimated last.

In order for the solution $(\hat{\beta}, \hat{\gamma}, \hat{\delta})$ to equations (39) to describe a relative maximum to $\ln L|n$, it is necessary and sufficient that the matrix of second partial derivatives ($\underline{\Sigma}$) be negative definite, see Frisch [1966, p. 120]. In examining this matrix in the case of (38), it is helpful to define $S(u,v) = \exp\{\beta u + \gamma v + \delta uv\}$. Then the function

$$s(u,v) = \frac{S(u,v)}{A}$$

has the properties:

(a) $s(u,v) \geq 0$

(b) $\int_0^{Y^*} \int_0^{X^*} s(u,v)\, dudv = 1$

(c) $s(u,v)$ is continuous on $[0 \leq u \leq X^*, 0 \leq v \leq Y^*]$.

Hence $s(u,v)$ is a probability density function [Gnedenko, 1962, p. 171].

Hence the matrix $\underline{\Sigma}$ can be shown to have diagonal elements such as

$$\sigma_{11} = -n\left[\int_0^{Y^*} \int_0^{X^*} u^2 s(u,v)\, dudv - \left(\int_0^{Y^*} \int_0^{X^*} us(u,v)\, dudv\right)^2\right]$$

$$= -n \operatorname{Var} U.$$

Continuing, the result is (where W is defined to be the W = UV)

$$\underline{\Sigma} = \begin{pmatrix} -n \text{ Var } U & -n \text{ Cov } (U,V) & -n \text{ Cov } (U,W) \\ -n \text{ Cov } (U,V) & -n \text{ Var } V & -n \text{ Cov } (V,W) \\ -n \text{ Cov } (U,W) & -n \text{ Cov } (V,W) & -n \text{ Var } W \end{pmatrix}$$

and $\underline{\Sigma}$ is revealed to be a covariance matrix. Note that the condition for a relative maximum, i.e. $\underline{\Sigma}$ negative definite, is independent of the realizations.

Now $\tilde{\underline{\Sigma}} = -n^3\underline{\Sigma}$ where $\tilde{\underline{\Sigma}}$ is the usual variance-covariance matrix for a tri-variate distribution. But $\tilde{\underline{\Sigma}}$ is positive semi-definite [Gnedenko, 1962, p. 212], hence $-\tilde{\underline{\Sigma}}$ is negative semi-definite. That each of the principal minors has non-zero determinants remains to be shown.

By the expressions given in Gnedenko [1966, p. 212], the covariance matrix $\tilde{\underline{\Sigma}}$ can be seen to be a Hankel matrix [Gantmacher, 1, 1959, p. 338]. Hence if the rows of $\tilde{\underline{\Sigma}}$ are linearly independent, then the determinant of $\tilde{\underline{\Sigma}} > 0$. But also Var U > 0 since U is a random variable and Var U Var V - $\text{Cov}^2(U,V) > 0$ since the case of line discontinuities has been excluded (i.e., U cannot be a linear function of V). By the same reasoning, W is linearly independent of U and V. Hence all principal minors are greater than zero, hence $\tilde{\underline{\Sigma}}$ is positive definite, hence $\underline{\Sigma}$ is negative definite. Thus $(\hat{\beta}, \hat{\gamma}, \hat{\delta})$ provides <u>at least</u> a relative maximum to ln L|n.

If it were possible to determine that $(\hat{\beta}, \hat{\gamma}, \hat{\delta})$ provides a global maximum to ln L|n in the region of interest, then conclusions as to uniqueness of the estimator could be drawn. Unfortunately, global extrema are difficult to establish. Since the method of estimation used was maximum likelihood, the estimates are consistent. Questions of biasedness are unresolved.

In order to solve the system of equations (39), it is necessary to determine initial values for the parameters as a starting point for an iterative procedure. The partial differentiation of ln L (35) with respect to the parameters and setting these partials equal to zero results, after some algebraic manipulation, in

$$n - \Lambda(X,Y) = 0$$

$$\Sigma x_i + \frac{\gamma}{\delta} - \frac{e^{\alpha}}{\delta}\left[\frac{e^{\gamma Y}(e^{(\beta+\delta Y)X} - 1)}{\beta+\delta Y} - \frac{e^{\beta X} - 1}{\beta}\right] = 0$$

$$\Sigma y_i + \frac{\beta}{\delta}\Lambda - \frac{e^{\alpha}}{\delta}\left[\frac{e^{\beta X}(e^{(\gamma+\delta X)Y} - 1)}{\gamma+\delta X} - \frac{e^{\gamma Y} - 1}{\beta}\right] = 0 \qquad (40)$$

$$\Sigma x_i y_i + \left(\frac{\beta\gamma+\delta}{\delta^2}\right)\Lambda + \frac{e^{\alpha}}{\delta(\beta+\delta Y)(\gamma+\delta X)}\left\{\frac{\beta\gamma}{\delta}\left[e^{\beta X+\gamma Y+\delta XY} - e^{\gamma Y} - e^{\beta X} - 1\right]\right.$$

$$\left. - \delta XY\left[e^{\beta X+\gamma Y+\delta XY} - 1\right] - \beta X\left[e^{\gamma Y} - 1\right] - \gamma Y\left[e^{\beta X} - 1\right]\right\} = 0$$

If it is assumed that the sum $\beta X + \gamma Y + \delta XY$ is small (near zero) as well as the individual terms in the summation being small, then the exponentials can be approximated by

$\exp\{x\} \cong 1 + x$, x near zero. Using the first equation in system (40) to give the value for $\Lambda(X^*,Y^*)$, i.e. $\Lambda(X^*,Y^*) = n$, and the linear approximation in the remaining terms gives the abbreviated system:

$$\begin{aligned}
&\Sigma x_i + \frac{\gamma}{\delta} n && = 0 \\
&\Sigma y_i + \frac{\beta}{\delta} n && = 0 \\
&\Sigma x_i y_i + \frac{\beta\gamma+\delta}{\delta^2} n - \frac{(1+\alpha)}{\delta(\beta+\delta Y)(\gamma+\delta X)} \left[\frac{2\beta\gamma}{\delta} + \beta\delta X^2 Y + \gamma\delta X Y^2 \right. \\
&\qquad \left. -\delta^2 XY + \beta\gamma XY\right] && = 0
\end{aligned} \tag{41}$$

The solution to (41) provides the initial estimates for the parameters. These estimates can then be used in (39) or (40) to search for sequentially closer and closer approximations in a mathematical programming approach.

Following the determination of the estimates $\hat{\beta}$, $\hat{\gamma}$ and $\hat{\delta}$, $\hat{\xi}$ can be determined from the solution to the firs equation in the set (40).

The determination of confidence intervals and levels of significance is not considered.

VI. CONCLUSIONS

The procedures in Chapters III-B, IV and V are dependent on the particular choice of parameter form; however, with different forms the concept of a non-homogeneous planar Poisson process may be used to describe a wide variety of "randomly" occurring phenomena. The choice of parameters which may be used is limited only by assumption I, i.e. positivity. One advantage of the method discussed herein over previously proposed schemes is the fact that the specific form used admits the possibility of a ridge or line of maximum density to be mathematically specified and estimated.

Also there is an attempt to describe the underlying process that caused the points to appear where they did, as opposed to using, for instance, the arc within which the most events were observed as the point estimate for the direction of maximum increase.

Further efforts in this area include a generalization into four dimensions (x,y,z,t) in order that zoological as well as botanical densities may be studied. Of especial interest is the estimation of densities of aquatic life and how the observed density fluctuates with season and with changes in environment. The latter problem seems of prime importance in evaluating the effects of anti-pollution programs on the fluid systems in which plants and animals exist.

Another problem which is closely related to the above is that of imperfect sampling and how the estimates are biased by sampling techniques.

Chapters III, IV and V may be redefined in terms of data gathered within a circle about some fixed point, especially with consideration of the relative efficiency of this data form referred to by Matern [1960].

APPENDIX A: THE BIVARIATE UNIFORM DISTRIBUTION

Given a region R in E^2 of area A and the fact that the probability of occurrence of an event in any sub-region R_1 of area A_1 within R is simply A_1/A, a bivariate uniform distribution is described. For definiteness assume the region R is rectangular, so $A = X^*Y^*$. Now

$$\text{Prob}\ (X \leq x, Y \leq y) = \frac{xy}{X^*Y^*} = \text{Prob}\ (X \leq x)\ \text{Prob}\ (Y \leq y)$$

for $0 \leq x \leq X^*$, $0 \leq y \leq Y^*$, in which case it is apparent that the coordinate axes define independently chosen univariate random variables.

Also, the density function is immediately

$$f(x,y) = 1/X^*Y^* \qquad 0 \leq x \leq X^*, \quad 0 \leq y \leq Y^*.$$

From the density function the joint density for n indepen ent bivariate uniform random variables is

$$f((x,y)_1,\ldots,(x,y)_n,n) = 1/(X^*Y^*)^n$$

where $(x,y)_i$ denotes the i^{th} pair of random variables selected. Now n pairs of random variables, or more simply n points in the plane, can only be ordered (without

replacement) in n! ways, independent of the ordering process chosen. Hence, the joint density function for n ordered bivariate uniform random variables is

$$f((x,y)_{(1)},\ldots,(x,y)_{(n)},n) = n!/(X^*Y^*)^n$$

where $(x,y)_{(i)}$ is the i^{th} point selected in the ordering scheme utilized.

As a specific example, consider the n points to be labelled with respect to increasing magnitude of the y-component. Then

$$y_k = y_{(k)} \quad k = 1,\ldots,n \quad \text{and} \quad (x,y)_{(k)} = (x_k,y_{(k)}).$$

If the x-components are also ordered, then the set of points $P = ((x_{(i)},y_{(j)});\ i,j = 1,\ldots,n)$ defines n^2 points, of which n are known to be "occupied," that is, to describe an event. For $x_{(1)}$, there exists some j such that $y_{(j)}$ gives the y-coordinate value for the event which gave rise to $x_{(1)}$. Similarly, for $x_{(2)}$ there are now n-1 j's remaining, one of which must correspond to the event giving rise to $x_{(2)}$. Continuing to $x_{(n)}$, there can only be one j left to be associated with the last x-value. Thus there are n! combinations of $(x,y)_{(i)}$, $i = 1,\ldots,n$ each having density of $1/(X^*Y^*)^n$ and so the ordered bivariate uniform density is established as that stated above.

LIST OF REFERENCES

1. Bartlett, M.S., "The spectral analysis of two-dimensional point processes," Biometrika, 51, p. 299, 1964.

2. Brown, M., "Statistical analysis of non-homogeneous Poisson processes," Stochastic Point Processes: Statistical Analysis, Theory and Application, edited by P.A.W. Lewis, Wiley-Interscience, New York, 1972.

3. Brownlee, K.A., Statistical Theory and Methodology in Science and Engineering, 2nd edition, Wiley, New York, 1965.

4. Cox, D.R. and Lewis, P.A.W., Statistical Analysis of Series of Events, Methuen and Co., Ltd., London, 1966.

5. Durbin, J., "Asymptotic Distributions of some Statistics Based on the Bivariate Sample Distribution Function," Non-parametric Techniques in Statistical Inference, edited by Madan Lal Puri, Cambridge Univ. Press, 1970.

6. Fisher, L., "A survey of the mathematical theory of multi-dimensional point processes," Stochastic Point Processes: Statistical Analysis, Theory and Applications, edited by P.A.W. Lewis, Wiley-Interscience, New York, 1972.

7. Frisch, R., Maxima and Minima: Theory and Economic Applications, Rand McNally, Chicago, 1966.

8. Gantmacher, F.R., The Theory of Matrices, 1, Chelsea Publishing Company, New York, 1959.

9. Gnedenko, B.V., The Theory of Probability, translated by B.D. Seckler, Chelsea, New York, 1962.

10. Holgate, P., "The use of distance methods for the analysis of spatial distributions of points," Stocnastic Point Processes: Statistical Analysis, Theory and Application, edited by P.A.W. Lewis, Wiley-Interscience, New York, 1972.

11. Jahnke, E. and Emde, F., Tables of Functions with Formulae and Curves, 4th edition, Dover, New York, 1945.

12. Kendall, M.G. and Stuart, A., The Advanced Theory of Statistics, 2, 3rd edition, Charles Griffin and Co., London, 1951.

13. Kowalski, C.J., "On the Effects of Non-normality on the Distribution of the Sample Product-moment Correlation Coefficient," Applied Statistics, 21, pp. 1-12, 1972.

14. Lamperti, John, Probability, W.A. Benjamin, Inc., New York, 1966.

15. Matern, B., Spatial Variation, Meddelanden Fran, Statens Skogsforskningsinstitut, Band 49, Nr 5, Stockholm, 1960.

16. Miles, R.E., "On the homogeneous planar Poisson point process," Mathematical Biosciences, 6, pp. 85-127, 1970.

17. Pearson, E.S. and Hartley, H.O., Biometrika Tables for Statisticians, 1, 3rd edition, Cambridge University Press, 1966.