This is to certify that the

dissertation entitled

SPATIAL PATTERNS: A STATISTICAL
FORMULATION AND ANALYSIS

presented by

Rangaswami Geetha

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics

Major professor

Date 11/5/81

SPATIAL PATTERNS: A STATISTICAL FORMULATION AND ANALYSIS

By

Rangaswami Geetha

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1981

ABSTRACT

SPATIAL PATTERNS: A STATISTICAL FORMULATION AND ANALYSIS

By

Rangaswami Geetha

In this dissertation a general study of spatial patterns is investigated. We have given a statistical formulation to the concept of spatial patterns, a problem which has long been overlooked by the ecologists, computer scientists etc. Have established the characterization of a random spatial pattern as a realization of a Poisson point process, through the notion of convergence of point processes. In the sequel we have introduced the stochastic integral with respect to a point process.

Further we have studied inferences on randomness (no spatial interaction) through estimation of the intensity of the spatial process and through testing hypothesis in a special subclass of spatial binary schemes described through near-neighbour systems. Here, we have established the chi-square behaviour of log-likelihood ratio under the null hypothesis of no spatial interaction without the use of Besag's (1974) coding method of estimation.

Finally, in our attempts to study the power of the test in this subclass of binary schemes, we established the contiguity of the probability measures under the specified hypotheses of interest but however realized

that the asymptotic distribution of the log-likelihood ratio under the alternative does not have an easy tractable form. A good conjecture is that it is a non-central chi-square distribution.

Dedicated to my mother and my late father

ACKNOWLEDGEMENT

I wish to thank Dr. V.S. Mandrekar for his guidance and encouragement during the preparation of this thesis. I would also like to thank Drs. J.C. Gardiner and D.C. Gilliland for their careful reading of the chapters and their valuable suggestions. My thanks are also due to Dr. J. Kinney for serving on my committee, to Michigan State University for the financial assistance and to Ms. Clara Hanna for typing of this thesis.

Finally, I sincerely appreciate the patience of my family members throughout my graduate study.

TABLE OF CONTENTS

CHAPTER 0

INTRODUCTION

The phrase spatial pattern is commonly used to describe the distribution of individuals in space. It is one of the topics investigated under the broader subject of pattern recognition otherwise known as the problem of classification to the statisticians.

The problem of classification has always concerned itself with classifying a sample of individuals into groups which are to be distinct in some sense. These groups may either be predetermined or determined using techniques of cluster analysis. However, in the world of organic nature, for example, with the distribution of plants and animals, the broad outlines of the spatial patterns are determined by the structural features of the physical environment. Therefore, in the study of spatial patterns it is not just enough to classify them into groups but rather determine whether or not the patterns exhibit randomness.

In Chapter I of this thesis we have given a statistical formulation of the above problem through the concept of point processes. It has been repeatedly maintained in the literature that by a random pattern is meant the pattern is a realization of a Poisson point process. Proposition 1.5.1 justifies this characterization. In the sequel, we need to introduce the notion of stochastic integral with respect to a point process and the notion of weak convergence for point processes (sec. 1.2 and sec. 1.4).

In Chapter II we discuss the ideas that are needed to construct valid spatial schemes through, what seems reasonable, near-neighbour systems. Section 2.1 introduces the notion of neighbours in a set of sites from a graph theoretic viewpoint [Berge - 1962]. Characterization of Markov random fields, Gibbs field and their equivalence for any finite graph are all discussed in section 2.2 [Carnal - 1979]. Theorem 2.2.1 often referred to as the Hammersley-Clifford theorem, plays a vital role in the construction of spatial schemes through near-neighbour systems. We use this theorem to characterize some of the specific spatial schemes (section 2.3).

In Chapter III, we look into some of the methods to determine randomness of a spatial pattern. Section 3.1 discusses some of the estimators of the intensity of the spatial process and their asymptotic properties. It also looks into the use of these estimators to study the randomness of the spatial process. In section 3.2 we consider a particular subclass of Markov random fields and give a method to test for randomness in this subclass. It has been shown that the log-likelihood ratio has a central chi-squared behaviour under the assumption of complete randomness of the spatial pattern. The proof involves simple use of Taylor's series expansion and follow the lines of proof of the classical theory on maximum likelihood estimation [Cramer - 1946].

In Chapter IV, we have attempted to discuss the power of the test against a specific alternative from the subclass of auto-binary schemes. The problem has been approached using ideas on contiguity -

a concept that describes the 'closeness' of sequences of probability measures. Section 4.1 discusses some of the characterizations of contiguity [Roussas - 1972]. In our formulation the classical techniques of contiguity fail. In section 4.2 using the basic principles of contiguity we establish the contiguity of probability measures under the null and alternative hypotheses. However, the asymptotic distribution of the log-likelihood ratio under the alternative does not seem to have a tractable form. A good conjecture is that the power of the test depends on a non-central chi-square distribution.

CHAPTER I

POINT PROCESSES AND SPATIAL PATTERNS

In this chapter we introduce the concept of a spatial pattern and the characterization of a random spatial pattern as a realization of a Poisson point process. We have approached the problem by introducing the notion of a point process and weak convergence of point processes. In the sequel we have defined stochastic integration with respect to a point process and Laplace functional of a point process.

1.1. Notion of a Point process:

Notations:

Let (Ω,F,P) be a probability space. Let $S = R^d$ be the d-dimensional Euclidean space and $\mathcal{B}(S)$ be the family of borel subsets of S. Let $x_1,x_2,\ldots$ denote points in S. Let A denote a compact set in S such that A contains finitely many x_i's.

Let

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise} \end{cases}$$

Let $M_+(S) = \{m$ on $\mathcal{B}(S) \mid \exists$ a countable set of points x_j such that $m(A) = \sum_j \delta_{x_j}(A) = \#$ of x-points in A, $A \in \mathcal{B}(S)$ and $m(A)$ is finite for all compact A in $\mathcal{B}(S)\}$.

Let $\mathcal{M}(S)$ denote the σ-algebra generated making all the mappings

$$m \to m(A) \qquad\qquad A \in \mathcal{B}(S)$$

of $M_+(S)$ into $Z^{+,\infty}$ (the set of non-negative integers including $+\infty$) measurable.

<u>Definition 1.1.1</u>: The family $\{\xi(A,\omega)\colon A \in \mathcal{B}(S)\}$ describes a point process if

a) $\xi(A,\omega)$ is non-negative integer valued $\forall\, A \in \mathcal{B}(S)$ and finite for compact A

b) $\xi(\cdot,\omega)$ is a measure such that $\xi(\cdot,\omega)$ puts mass 0 or 1 on singleton sets i.e. (i) For a sequence $A_1, A_2, \ldots, A_m$ of pairwise disjoint sets in $\mathcal{B}(S)$ we have

$$\xi\left(\bigcup_{j=1}^{m} A_j,\omega\right) = \sum_{j=1}^{m} \xi(A_j,\omega)$$

(ii) For $A_1 \supset A_2 \supset \ldots$ in $\mathcal{B}(S)$ such that $\bigcap_n A_n = \phi$ we have

$$\lim_n \xi(A_n,\omega) = 0$$

c) $\xi(A,\cdot)$ is measurable on Ω into $Z^{+,\infty}$.

<u>Definition 1.1.2</u>: A point process ξ on S is a measurable map from (Ω,F,P) into $(M_+(S),\ \mathcal{M}(S))$ i.e., $\forall\, \omega \in \Omega\ \xi(\omega)$ is an $(M_+(S),\ \mathcal{M}(S))$-valued random variable.

<u>Theorem 1.1.1</u>: The above two notions of a point process are equivalent.

<u>Proof</u>: We first prove the easy part: Assume that we are given an $(M_+(S),\ \mathcal{M}(S))$ - valued random variable ξ defined on (Ω,F,P). For every $A \in \mathcal{B}(S)$ and $\omega \in \Omega$, define

$$\xi(A,\omega) = \xi(\omega)(A)$$

Then $\xi(\cdot,\omega)$ satisfies all the conditions a), b) and c) of definition 1.1.1.

The converse implication of the theorem depends upon a Kolmogorov type theorem. The proof of this part is given in Appendix A.

Remark 1.1.1: Following theorem 1.1.1, we shall find it convenient to think of a point process sometimes as a measurable map on (Ω,F,P) and sometimes as a set function satisfying a), b) and c) of definition 1.1.1.

Definition 1.1.3: The intensity of a point process ξ: $(\Omega,F,P) \to (M_+(S), \mathcal{M}(S))$ is defined as the measure λ on S such that $\forall\ A \in \mathcal{B}(S)$

$$\lambda(A) = E_P\{\xi(A,\omega)\}$$

i.e.

(1.1.1) $$\lambda(A) = \int_\Omega \xi(A,\omega)P(d\omega)$$

Example 1.1.1: Poisson point process

A point process $\{\xi(A,\omega)\colon\ A \in \mathcal{B}(S)\}$ is called a Poisson point process with intensity $\lambda > 0$ if

(i) $\forall\ A \in \mathcal{B}(S)$, $\xi(A,\omega)$ (which by definition 1.1.2 indicates the number of points in A) has a Poisson distribution with parameter $\lambda(A)$ (if $\lambda(A) = +\infty$, $\xi(A,\omega) = +\infty$);

(ii) $\forall$ finite collection $\{A_1,\ldots,A_m\}$ of disjoint sets in $\mathcal{B}(S)$, the random variables $\xi(A_1,\cdot),\ldots,\xi(A_m,\cdot)$ are mutually independent.

To show that the above formulation characterizes a point process we need to show that the function

$$q(A_1,\ldots,A_m;\ r_1,\ldots,r_m) = \prod_{j=1}^{m}\left\{e^{-\lambda(A_j)}\,\frac{[\lambda(A_j)]^{r_j}}{r_j!}\right\}$$

satisfies conditions A-1(i-iv) of the Appendix A, where $\{A_j:\ j = 1,\ldots,m\}$ is a sequence of disjoint sets and λ is the intensity of the process.

<u>A-1(i)</u> $q(A_1,\ldots,A_m;\ r_1,\ldots r_m)$ is obviously a probability distribution on the m-tuples of non-negative integers $r_1,\ldots,r_m$ and

$$q(A_1,A_2;\ r_1,r_2) = q(A_2,A_1;\ r_2,r_1)$$

<u>A-1(ii)</u> The functions q are consistent

i.e., $$\sum_{r_2=0}^{\infty} q(A_1,A_2;\ r_1,r_2) = q(A_1,r_1)$$

$$\text{Lhs} = \sum_{r_2=0}^{\infty} q(A_1,A_2;\ r_1,r_2)$$

$$= \sum_{r_2=0}^{\infty} e^{-(\lambda(A_1) + \lambda(A_2))}\,\frac{(\lambda(A_1))^{r_1}\lambda(A_2)^{r_2}}{r_1!\ r_2!}$$

$$= \frac{e^{-\lambda(A_1)}(\lambda(A_1))^{r_1}}{r_1!}\quad e^{-\lambda(A_2)}\sum_{r_2=0}^{\infty}\frac{(\lambda(A_2))^{r_2}}{r_2!}$$

$$= e^{-\lambda(A_1)}\,\frac{[\lambda(A_1)]^{r_1}}{r_1!} = q(A_1,r_1)$$

A-1(iii) Let $A_1,\dots,A_m$ be disjoint sets such that $A = A_1 \cup \dots \cup A_m$. Then following remark A-1, (A-1-1) implies that

$$q(A,A_1,A_2,\dots,A_m;\ r,r_1,\dots,r_m) = 0$$

unless $r = r_1 + \dots + r_m$ and

$$q(A,A_1,\dots,A_m,r_1 + \dots + r_m,\ r_1,\dots,r_m)$$

$$= q(A_1,\dots,A_m;\ r_1,\dots,r_m).$$

A-1(iv) Let $A_1 \supset A_2 \supset \dots$ be such that $\bigcap_n A_n = \phi$ which implies $\lambda(A_n) \to 0$. Consequently $\lim_n q(A_n,0) = \lim_n e^{-\lambda(A_n)} = 1$. Thus the formulation of a Poisson process characterizes a point process.

Remark 1.1.2: In example 1.1.1 if $\lambda(A) = c \,.\, \nu(A)$ where c is a constant and ν is the Lebesgue measure then the Poisson process is known as a homogeneous Poisson process with intensity c.

1.2. Stochastic Integral with respect to a Point process:

Notations:

Let $C_k^+(S)(C_{ks}^+(S))$ denote the family of all non-negative continuous functions (simple functions) on S with a compact support.

Let ξ be a point process with intensity λ.

Definition 1.2.1: Let $f \in C_{ks}^+(S)$ so that

$$f(u) = \sum_{j=1}^{m} c_j I_{A_j}(u)$$

where c_j's are positive constants and I_{A_j}'s are the indicator functions of the disjoint sets $A_1,\ldots,A_m$.

Then the stochastic integral of f with respect to ξ is defined by

$$\int_S f(u)\xi(du,\omega) = \sum_{j=1}^{m} c_j\ \xi(A_j,\omega) \qquad \forall\ \omega \in \Omega \tag{1.2.1}$$

Note that in the above definition if A is a compact support of f, then

$$\xi(A,\omega) < \infty \qquad \forall\ \omega \text{ (def. 1.1.1)}$$

which implies by monotonicity of $\xi(\cdot,\omega)$ (definition 1.1.1(b)) that

$$\xi(A_j,\omega) < \infty \qquad \forall j = 1,2,\ldots,m;\ \forall\ \omega \in \Omega$$

Consequently it follows that

$$\int_S f(u)\xi(du,\omega) < \infty \qquad \forall\ \omega \in \Omega.$$

Remark 1.2.1: Let $f \in C^+_{ks}(S)$

$$\begin{aligned} E_P\{\int_S f(u)\xi(du,\omega)\} &= \int_\Omega \{\int_S f(u)\xi(du,\omega)\}\ P(d\omega) \\ &= \int_S f(u) \int_\Omega \xi(du,\omega)P(d\omega) \quad \text{(Fubini's theorem)} \\ &= \int_S f(u)\lambda(du) \qquad\qquad \text{(by 1.1.1)} \end{aligned}$$

thus, $\forall\ f \in C^+_{ks}(S)$ we have

$$E_P\{\int_S f(u)\xi(du,\omega)\} = \int_S f(u)\lambda(du). \tag{1.2.2}$$

Let now $f \in L_1^+(\lambda)$, then there exists a sequence $\{f_n\}$ of simple functions in $L_1^+(\lambda)$ such that

$$\|f_k - f_j\|_1 \to 0 \qquad k,j \text{ integers, } k \geq j$$

Now

$$|E_P\{\int_S f_k(u)\xi(du,\omega) - \int_S f_j(u)\xi(du,\omega)\}|$$

$$\leq E_P \int_S |f_k(u) - f_j(u)| \; \xi(du,\omega)$$

$$= \int_S |f_k(u) - f_j(u)| \; \lambda(du) \qquad \text{(by 1.2.2)}$$

$$= \|f_k - f_j\|_1$$

$$\to 0$$

Hence it follows that $\int_S f_n(u)\xi(du,\omega)$ converges in $L_1(P)$

We denote this L_1 - limit by

$$\int_S f(u)\xi(du,\omega).$$

<u>Definition 1.2.2</u>: The stochastic integral of an f in $L_1^+(\lambda)$ is defined by

$$\int_S f(u)\xi(du,\omega) = L_1 - \lim_n \int_S f_n(u)\xi(du,\omega) \tag{1.2.3}$$

where $f_n \in L_1^+(\lambda)$ are simple.

<u>Remark 1.2.2</u>: The above definition 1.2.2 is independent of the particular sequence $\{f_n\}$. For if $\{g_n\}$ is another sequence of simple functions in $L_1^+(\lambda)$ converging to f in the sense that

$$\int_S f(u)\xi(du,\omega) = L_1\text{-}\lim_n \int_S g_n(u)\xi(du,\omega)$$

then the sequence $\{h_n\}$ where $h_{2n} = f_n$ and $h_{2n+1} = g_n$ is also convergent to f in the same sense.

i.e.,

$$\int_S f(u)\xi(du,\omega) = L_1\text{-}\lim_n \int_S h_n(u)\xi(du,\omega).$$

Consequently it follows that

$$L_1\text{-}\lim_n \int_S f_n(u)\xi(du,\omega) = L_1\text{-}\lim_n \int_S g_n(u)\xi(du,\omega) \qquad \text{a.s.}$$

Proposition 1.2.1: Let f_1, f_2 be functions in $L_1^+(\lambda)$ and let a_1, a_2 be real numbers so that $a_1f_1 + a_2f_2$ is in $L_1^+(\lambda)$. Also then

$$\int_S [a_1f_1(u) + a_2f_2(u)]\ \xi(du,\omega)$$

$$= a_1 \int_S f_1(u)\xi(du,\omega) + a_2 \int_S f_2(u)\xi(du,\omega)$$

Proof: Obvious.

Remark 1.2.3: Let $f \in C_k^+(S)$ so that there exists a sequence $\{f_n\}$ in $C_{ks}^+(S)$ such that

$$f = \lim_n f_n$$

Since $C_k^+(S) \subset L_1^+(\lambda)$ we can define the stochastic integral of an $f \in C_k^+(S)$ by (1.2.3).

Consequently (1.2.2) is true for any $f \in C_k^+(S)$

i.e., $\forall\ f \in C_k^+(S)$ we have

(1.2.4) $$E_P\{\int_S f(u)\xi(du,\omega)\} = \int_S f(u)\lambda(du).$$

1.3 Laplace functional of a point process:

Definition 1.3.1: The Laplace transform of a probability measure Q on the space $(M_+(S),\ \mathcal{M}(S))$ is defined by

$$(1.3.1)\qquad \Psi_Q(f) = \int_{M_+(S)} \exp\{-\int_S f(u)m(du)\}Q(dm) \qquad \forall\ f \in C_k^+(S).$$

Definition 1.3.2: The Laplace functional of a point process ξ is defined to be the Laplace transform of its probability law P_ξ

i.e., $\forall\ f \in C_k^+(S)$

$$\Psi_\xi(f) = \int_{M_+(S)} \exp\{-\int_S f(u)m(du)\}P_\xi(dm)$$

By transformation of variables, this gives

$$(1.3.2)\qquad \Psi_\xi(f) = \int_\Omega \exp\{-\int_S f(u)\xi(du,\omega)\}P(d\omega) \qquad \forall\ f \in C_k^+(S)$$

Lemma 1.3.1: For every increasing sequence $\{f_n\}$ of functions in $C_k^+(S)$ we have

$$(1.3.3)\qquad \Psi_\xi(\lim_n \uparrow f_n) = \lim_n \downarrow \Psi_\xi(f_n).$$

Proof: Let $f = \lim_n f_n$.
By (1.3.2) we have

$$\Psi_\xi(f_n) = \int_\Omega \exp\{-\int_S f_n(u)\xi(du,\omega)\}P(d\omega)$$

and

$$\Psi_\xi(f) = \int_\Omega \exp\{-\int_S f(u)\xi(du,\omega)\}P(d\omega)$$

By (1.2.3)

$$\int_S f(u)\xi(du,\omega) = L_1\text{-}\lim_n \int_S f_n(u)\xi(du,\omega)$$

implies

$$\int_S f_n(u)\xi(du,\omega) \xrightarrow{\text{Prob}} \int_S f(u)\xi(du,\omega)$$

implies

$$\exp\{-\int_S f_n(u)\xi(du,\omega)\} \xrightarrow{\text{Prob}} \exp\{-\int_S f(u)\xi(du,\omega)\}.$$

Also,

$\exp\{-\int_S f_n(u)\xi(du,\omega)\}$ are bounded by 1. Thus by Lebesgue dominated convergence theorem it follows that

$$\Psi_\xi(f) = \lim_n \Psi_\xi(f_n)$$

as was to be proved.

<u>Example 1.3.1</u>: The Laplace functional of a Poisson point process ξ with intensity λ is given by

$$(1.3.4) \qquad \Psi_\xi(f) = \exp\{-\int_S [1-e^{-f(u)}]\lambda(du)\} \qquad \forall\ f \in C_k^+(S)$$

And conversely, a point process ξ whose Laplace functional is of the form given by (1.3.4) $\forall\ f \in C_k^+(S)$, is a Poisson point process with intensity λ.

<u>Proof</u>: By lemma 1.3.1 it is enough to consider functions f of the form

$$f(u) = c_1 I_{A_1}(u) + \ldots + c_m I_{A_m}(u)$$

where $c_1,\ldots,c_m$ are positive constants and $I_{A_1},\ldots,I_{A_m}$ are indicator functions of a set of disjoint sets $A_1,\ldots,A_m$ respectively. Following (1.3.2) the Laplace functional of a point process ξ is given by

$$\Psi_\xi(f) = \int_\Omega \exp\{-\int_S f(u)\xi(du,\omega)\}P(d\omega)$$

Here ξ is a Poisson point process with intensity λ so that P is a probability measure determined by functions q satisfying:

$$q(A_1,\ldots,A_m;\ r_1,\ldots,r_m) = \prod_{j=1}^{m} \{\frac{e^{-\lambda(A_j)}(\lambda(A_j))^{r_j}}{r_j!}\}$$

where $r_1,\ldots,r_m$ are non-negative integers and A_j's are disjoint sets. Consequently it follows that

$$\begin{aligned}
\Psi_\xi(f) &= \int_\Omega \exp\{-\sum_{j=1}^{m} c_j\ \xi(A_j,\omega)\}P(d\omega) \\
&= \sum_{r_1,r_2,\ldots} \{e^{-\sum c_j r_j}\ e^{-\sum\lambda(A_j)} \prod_{j=1}^{m} \frac{[\lambda(A_j)]^{r_j}}{r_j!}\} \\
&= \exp\{-\sum_j \lambda(A_j)\}\exp\{\lambda(A_1)e^{-c_1} + \ldots + \lambda(A_m)e^{-c_m}\} \\
&= \exp\{-\sum_j \lambda(A_j)(1 - e^{-c_j})\} \\
&= \exp\{-\int_S [1 - e^{-f(u)}]\ \lambda(du)\}.
\end{aligned}$$

Thus, by lemma (1.3.1) the Laplace functional of a Poisson point process ξ with intensity λ is given by

$$\exp\{-\int_S [1 - e^{-f(u)}]\ \lambda(du)\} \qquad \forall\ f \in C_k^+(S)$$

Conversely: let $A_1,\dots,A_m$ be a finite collection of disjoint sets in $\mathcal{B}(S)$. Then the generating function of the random vector $(\xi(A_1,\cdot),\dots,\xi(A_m,\cdot))$ is given by

$$E(s_1^{\xi(A_1)} \dots s_m^{\xi(A_m)}) \quad 0 < s_i \leq 1,\ i = 1,\dots,m$$

$$= E\{\exp[-\sum_1^m \xi(A_i,\cdot) \log \frac{1}{s_i}\}$$

$$= E \exp\{-\int_S f(u)\xi(du,\omega)\}$$

$$\text{(for } f(u) = \sum_1^m \log \frac{1}{s_i} I_{A_i}(u) \in C_k^+(S))$$

$$= \Psi_\xi(f)$$

$$= \exp\{-\int_S [1 - e^{-f(u)}]\ \lambda(du)\} \text{ (by given hypothesis)}$$

$$= \exp\{-\sum_1^m (1 - s_i)\ \lambda(A_i)\}$$

which is the generating function of a product of m independent Poisson random variables with parameters $\lambda(A_1),\dots,\lambda(A_m)$ respectively.

Since generating functions determine a distribution uniquely it follows that a point process ξ with intensity λ whose Laplace functional is given by (1.3.4) must be a Poisson point process with intensity λ.

1.4 Notion of weak convergence for point processes:

Definition 1.4.1: A sequence $\{\lambda_n\}$ of measures on $(S,\mathcal{B}(S))$ is said to converge weakly to a measure λ on $(S,\mathcal{B}(S))$ if

$$(1.4.1) \qquad \int_S f(u)\lambda_n(du) \rightarrow \int_S f(u)\lambda(du)$$

for every continuous, bounded real valued function f on S.

Theorem 1.4.1: [Neveu - 1976, p. 282]

A sequence $\{P_n\}$ of probability measures on $(M_+(S), \mathcal{M}(S))$ converges weakly to a probability measure P on the same space if and only if

$$(1.4.2) \qquad \Psi_{P_n}(f) \rightarrow \Psi_P(f) \quad \forall f \in C_k^+(S)$$

Proposition 1.4.1: Let $\{\xi_n(A,\omega): A \in \mathcal{B}(S)\}$ be a sequence of point processes with intensity $\{\lambda_n\}$ and $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$ be a point process with intensity λ.

Then: $\lambda_n \rightarrow \lambda$ weakly implies $\forall f \in C_k^+(S)$:

$$(1.4.3) \quad (i) \qquad \int_S f(u)\xi_n(du,\cdot) \xrightarrow{L_1} \int_S f(u)\xi(du,\cdot)$$

and

$$(1.4.4) \quad (ii) \qquad \Psi_{\xi_n}(f) \rightarrow \Psi_\xi(f).$$

Proof: (i) Let $f \in C_k^+(S)$.

Now

$$|E_P\{\int_S f(u)\xi_n(du,\cdot) - \int_S f(u)\xi(du,\cdot)\}|$$

$$= |\int_S f(u)\lambda_n(du) - \int_S f(u)\lambda(du)| \qquad \text{(by (1.2.4))}$$

$$\rightarrow 0 \quad \text{(since by hypothesis } \lambda_n \rightarrow \lambda \text{ weakly).}$$

Thus it follows that $\forall\ f \in C_k^+(S)$ $\int_S f(u)\xi_n(du,\cdot)$ converges to $\int_S f(u)\xi(du,\cdot)$ in $L_1(P)$. This establishes (1.4.3).

(ii) (1.4.3) implies

$$\int_S f(u)\xi_n(du,\cdot) \xrightarrow{\text{Prob}} \int_S f(u)\xi(du,\cdot)$$

implies

$$\exp\{-\int_S f(u)\xi_n(du,\cdot)\} \xrightarrow{\text{Prob}} \exp\{-\int_S f(u)\xi(du,\cdot)\}$$

and $\exp\{-\int_S f(u)\xi_n(du,\cdot)\}$ are bounded by 1. Thus by Lebesgue dominated convergence theorem and using (1.3.2) it follows that

$$\Psi_{\xi_n}(f) \to \Psi_{\xi}(f) \qquad \forall\ f \in C_k^+(S)$$

thereby establishing (1.4.4).

<u>Definition 1.4.2</u>: A sequence $\{\xi_n(A,\omega)\colon\ A \in \mathcal{B}(S)\}$ of point processes is said to converge in distribution to a point process $\{\xi(A,\omega)\colon A \in \mathcal{B}(S)\}$ if and only if

$$P_{\xi_n} \text{ converges weakly to } P_{\xi}$$

where P_{ξ_n} and P_{ξ} are respectively the probability laws of ξ_n and ξ.

We prove in theorem 1.4.2, that the weak convergence of point processes is guaranteed by the convergence of the corresponding finite dimensional distributions (definition A-1 of Appendix A).

Theorem 1.4.2: For $\{\xi_n(A,\omega)\colon A \in \mathcal{B}(S)\}$ of point processes and a point process $\{\xi(A,\omega)\colon A \in \mathcal{B}(S)\}$ the following three statements are equivalent:

(i) P_{ξ_n} converges weakly to P_ξ

(ii) $\Psi_{\xi_n}(f) \to \Psi_\xi(f) \quad \forall f \in C_k^+(S)$

(iii) corresponding finite dimensional distributions converge weakly.

i.e. $\text{Prob}\,\{\xi_n(A_1,\omega) = r_1,\ldots,\xi_n(A_m,\omega) = r_m\}$

$$\longrightarrow \text{Prob}\,\{\xi(A_1,\omega) = r_1,\ldots,\xi(A_m,\omega) = r_m\}$$

for $A_1,\ldots,A_m$ in $\mathcal{B}(S)$ and $r_1,\ldots,r_m$ nonnegative integers.

Proof: 1°: (i) $\Leftrightarrow$ (ii) (see theorem 1.4.1)

2°: we will prove (ii) $\Leftrightarrow$ (iii)

a) (ii) $\Rightarrow$ (i) (by 1° above)

and (i) $\Rightarrow$ (iii) [Billingsley - 1968]

thus (ii) $\Rightarrow$ (iii).

b) (iii) $\Rightarrow$ (ii): For

By lemma (1.3.1) it is enough to consider functions f in $C_{ks}^+(S)$ i.e. of the form

$$f(u) = \sum_{j=1}^{m} c_j\, I_{A_j}(u)$$

where c_j's are positive constants and I_{A_j}'s are the indicator functions of disjoint sets A_j $(j = 1,2,\ldots,m)$.

By (1.3.2) we have $\forall f \in C_{ks}^+(S)$:

$$\Psi_{\xi_n}(f) = \int_\Omega \exp\{-\int_S f(u)\xi_n(du,\omega)\ P(d\omega)$$

$$= \int_\Omega \exp\{-\sum_j c_j\ \xi_n(A_j,\omega)\}P(d\omega)$$

$$= \sum_{r_1, r_2 \ldots} \exp\{-\sum_j c_j r_j\} \text{ Prob } \{\xi_n(A_1,\omega) = r_1, \ldots \xi_n(A_m,\omega) = r_m\}$$

$$\to \sum_{r_1, r_2, \ldots} \exp\{-\sum_j c_j r_j\} \text{ Prob } \{\xi(A_1,\omega) = r_1, \ldots, \xi(A_m,\omega) = r_m\}$$

(by given hypothesis (iii))

$$= \int_\Omega \exp\{-\int_S f(u)\xi(du,\omega)\}P(d\omega)$$

$$= \Psi_\xi(f)$$

i.e. $\Psi_{\xi_n}(f)$ converges to $\Psi_\xi(f)$ $\forall$ $f \in C^+_{ks}(S)$. Thus it follows from lemma 1.3.1, that

$$\Psi_{\xi_n}(f) \to \Psi_\xi(f) \qquad \forall\ f \in C^+_k(S)$$

1^o and 2^o together imply the theorem.

Remark 1.4.1: By theorem 1.4.2 any one of the three equivalent statements imply the convergence in distribution of the sequence $\{\xi_n(A,\omega): A \in \mathcal{B}(S)\}$ of point processes to a point process $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$.

1.5 Concept of spatial patterns:

Spatial pattern is most commonly used by plant ecologists to describe the distribution of plants in a given area of study. In a more generality, by a spatial pattern is meant the distribtuion of points in space i.e., a spatial pattern is nothing but a realization of a point process.

Study of spatial patterns is encountered widely in the areas such as ecology, geology, medicine, forestry, image processing etc. Some of the specific examples of spatial patterns include

(i) distribution of stars in a galaxy;

(ii) distribution of a number of trees in a forest;

(iii) patterns of various rock formations on a geologic map;

(iv) texture modelling (through the description of images);

(v) epidemic spread from the map of a city.

The following proposition characterizes the behaviour of point distributed completely randomly.

<u>Proposition 1.5.1</u>: Suppose that there are N_M individuals distributed uniformly over a region M of area $|M|$ such that

$$(1.5.1) \qquad \frac{N_M}{|M|} \to \text{ a constant (say } c)$$

then as the region of study is expanded into the plane i.e. as $|M| \to \infty$ the distribution of events approaches the Poisson distribution of events in the sense that:

If $A_1,\ldots,A_m$ are any disjoint sets and $\xi(A_1),\ldots,\xi(A_m)$ denote respectively the number of individuals in $A_1,\ldots,A_m$ then $\xi(A_i)$: $i = 1,\ldots,m$ are random variables such that

$$\lim_{|M|\to\infty} \text{Prob}\{\xi(A_i) = r_i;\ r = 1,\ldots,m\} = \prod_{i=1}^{m} \{e^{-c\nu(A_i)} \frac{(c\nu(A_i))^{r_i}}{r_i!}$$

ν being the Lebesgue measure.

Proof: Let $\nu(A) = \sum_{i=1}^{m} \nu(A_i)$, $r = \sum_{i=1}^{m} r_i$.

Now

$$\text{Prob}\{\xi(A_i) = r_i;\ i = 1,\dots,m\} = \frac{N_M!}{r_1!\dots r_m!(N_M-r)!} \prod_{i=1}^{m} \left(\frac{\nu(A_i)}{|M|}\right)^{r_i} \times \left(\frac{|M|-\nu(A)}{|M|}\right)^{N_M-r}$$

$$= \prod_{i=1}^{m} \frac{(\nu(A_i))^{r_i}}{r_i!} \; \frac{N_M!}{(N_M-r)!\ |M|^r} \; \left(1-\frac{\nu(A)}{|M|}\right)^{N_M-r}$$

$$= \prod_{i=1}^{m} \frac{(\nu(A_i))^{r_i}}{r_i!} \left(1-\frac{\nu(A)}{|M|}\right)^{N_M-r} \frac{N_M}{|M|} \cdots \frac{N_M-r+1}{|M|}$$

$$\xrightarrow{|M|\to\infty} \prod_{i=1}^{m} \frac{(\nu(A_i))^{r_i}}{r_i!} \; e^{-c\nu(A)} \; c^r$$

$$= \prod_{i=1}^{m} \frac{(c\nu(A_i))^{r_i}}{r_i!} \; e^{-c\nu(A_i)}$$

as was to be proved.

Remark 1.5.1: Limit in equation (1.5.1) is referred to as thermodynamic limit.

Remark 1.5.2: Proposition 1.5.1 says that provided $\frac{N_M}{|M|} \to$ a constant then the finite dimensional distributions of a uniformly distributed point process approaches the finite dimensional distributions of a homogeneous Poisson point process. Consequently by theorem 1.4.2 it follows that in the thermodynamic limit sense, a uniformly distributed point process $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$ approaches a homogeneous Poisson point process. Thus following proposition 1.5.1 a spatial pattern $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$ is called random if it is a homogeneous Poisson point process. In that case we also say that $\xi(\cdot,\cdot)$ is a realization of a homogeneous Poisson point process.

Remark 1.5.3: For a random pattern that is a realization of a homogeneous Poisson point process with intensity λ, λ is also referred to as the expected number of individuals per unit area.

The following remark comments on the types of spatial schemes that we shall or shall not be discussing.

Remark 1.5.4: At a formal level, we shall largely be concerned with a rather arbitrary system. Later on we shall confine our attention to specific spatial schemes. However, it is to be realized that a spatial scheme could have developed continuously through time. For example: incidence of spotted wilt over a rectangular array of tomato plants. The disease is passed on by insects and after a period of time we could expect to observe clusters of infected plants. Such spatial schemes are referred to as spatial temporal schemes. We shall not be dealing with such schemes in this study. Thus, we shall only be dealing with homogeneous spatial schemes at an isolated instant of time. In many practical situations, this is reasonable since we can only observe the variables at a single point in time.

CHAPTER II

MARKOV RANDOM FIELDS

In Chapter I we introduced the notion of a spatial pattern and showed why a random pattern can be considered as a realization of a homogeneous Poisson point process. In this chapter we discuss the notion of neighbours in a set of sites from a graph theoretic viewpoint. Further, we give a characterization of Markov random field and Gibbs field and present a proof of the theorem which gives the equivalence of Markov random field and Gibbs field with near-neighbour potential for any finite graph.

Finally we have shown that the representing measure for a Markov random field with no interaction is the Poisson measure so that under the assumption of randomness spatial schemes have an exponential density indexed by V. Grimmett's potential, with respect to the Poisson measure.

2.1 Notion of neighbours in a set of sites:

Before we describe the concept of neighbours for a set of sites we need a few definitions from the theory of graphs [Berge - 1962].

Definiton 2.1.1: A graph denoted by $G = (\Lambda,\Gamma)$ is the pair consisting of a set Λ and a function Γ mapping Λ into Λ.

Here Λ is known as the set of vertices (or sites) of the graph G. The graph G is called finite if Λ has a finite number of elements.

For $s \in \Lambda$, let Γs denote the image of s under Γ. Consequently, if $A \subset \Lambda$ then the image of A under Γ is the set

$$\Gamma A = \bigcup_{s \in A} \Gamma s.$$

<u>Definition 2.1.2</u>: The pair (s,t) with $t \in \Gamma s$ is called an arc of the graph G.

Let U denote the set of all arcs of the graph G. We shall use (Λ,Γ) or (Λ,U) interchangeably to represent a graph G.

<u>Definition 2.1.3</u>: Path is a sequence $(u_1,u_2,...)$ of arcs of a graph G such that the terminal vertex of each arc coincides with the initial vertex of the succeeding arc.

If a path γ meets in turn the vertices $s_1,...,s_k$ one may also write

$$\gamma = [s_1,...,s_k].$$

The length of a path $\gamma = (u_1,...,u_k)$ is the number of arcs in the sequence (say $\ell(\gamma) = k$).

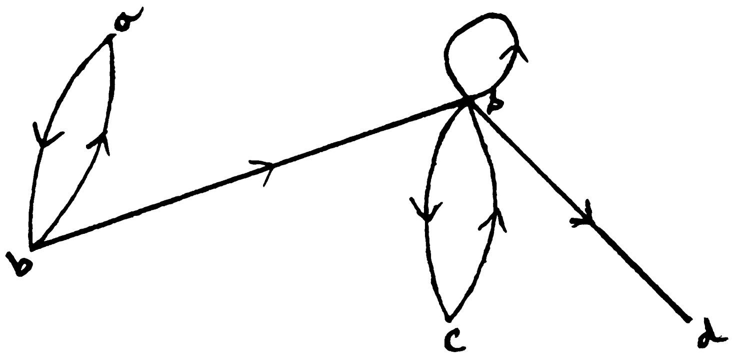
<u>Definition 2.1.4</u>: A circuit is a finite path in which the initial vertex coincides with the terminal vertex, and a loop is a circuit of length 1, consisting of the single arc (s,s).

<u>Definition 2.1.5</u>: In a graph $G = (\Lambda,U)$ an edge is a set of two elements $s,t \in \Lambda$ such that

$$\text{either } (s,t) \in U \text{ or } (t,s) \in U.$$

Note that the concept of an edge should not be confused with that of an arc which implies an orientation.

Example 2.1.1: Consider the following figure:



Here U is the set of arcs (a,b), (b,a) (b,s), (s,s), (s,c), (c,s) and (s,d).

The sequence $\{(a,b), (b,s), (s,c)\}$ represents a path of length 3. The path $\{(s,c), (c,s)\}$ is a circuit and (s,s) is a loop.

Finally, there are five edges namely (a,b), (b,s), (s,c), (s,d) and (s,s).

With the above notations and definitions we are now ready to describe the notion of neighbours in a set of sites.

Defintion 2.1.6: Let $G = (\Lambda,\Gamma)$ be any undirected finite graph with no loops or multiple edges. Then two sites $s,t \in \Lambda$ are called neighbours if there is an edge between s and t.

Further, two sites $s,t \in \Lambda$ are called neighbours of rth order if there are at most r edges between s and t.

Remark 2.1.1: (i) Neighbours of order 1 are also referred to as nearest neighbours;

(ii) The above definition of neighbours does not imply that the neighbours of a site are necessarily close in terms of distances.

Example 2.1.2: Consider a rectangular lattice with points labelled with integer pairs (i,j). Then figure 1 represents neighbours of order 1 and figure 2 represents neighbours of order 2, of an arbitrary point (i,j).

	$(i,j+1)$	
$(i-1,j)$	X (i,j)	$(i+1,j)$
	$(i,j-1)$	

Figure (1)

$(i-1,j+1)$	$(i,j+1)$	$(i+1,j+1)$
$(i-1,j)$	X (i,j)	$(i+1,j)$
$(i-1,j-1)$	$(i,j-1)$	$(i+1,j-1)$

Figure (2)

2.2 Definitions and characterizations of Markovian fields and Gibbs fields:

The following exposition follows Carnal (1979).

Notations: In what follows:

Let $G = (\Lambda,\Gamma)$ be a finite graph with $|\Lambda| = n$.

Let $\{X_t: t \in \Lambda\}$ denote a family of random variables taking values in some measure space $(E,\mathcal{E},\mu)$ where μ is a σ-finite measure.

Let $\varphi \in E^\Lambda$ i.e. $\varphi: \Lambda \to E$. For $A \subset \Lambda$, let ∂A denote the set of neighbours of sites in A.

Let $\bar{A} = A \cup \partial A$ and A^c denote the complement of A; ϕ denote the empty set.

Let $\varphi_A = \varphi|_A$ denote the restriction of φ to A.
Further, if $\theta \in E$

$$\theta_{L(\varphi)} = \{t \in \Lambda \mid \varphi(t) \neq \theta\}$$

$$\theta_\varphi^A(t) = \begin{cases} \varphi(t) & \text{if } t \in A \\ 0 & \text{otherwise.} \end{cases}$$

Also when no confusion is involved we shall write ∂t for $\partial\{t\}$, the set of neighbours of site t and $\varphi_t = \varphi_{\{t\}} = \varphi(t)$.

Definition 2.2.1: We say that $\{X_t:\ t \in \Lambda\}$ is a Markovian field on Λ if

(i) there exists an $f > 0$ on $E\times\ldots\times E$ (n times) such that

(2.2.1) $$P[X_{t_1} \in A_1,\ldots,X_{t_n} \in A_n] = \int_{A_1\times\ldots\times A_n} f \, d \otimes_n \mu \qquad \otimes_n \mu\text{-a.s.}$$

where $A_1,\ldots,A_n \in E$ and $\Lambda = \{t_1,\ldots,t_n\}$; and

(2.2.2) (ii) $$f(\varphi_A|\varphi_{A^c}) = f(\varphi_A|\varphi_{\partial A})$$

for $A \subset \Lambda$, $\varphi \in E^\Lambda$ where

(2.2.3) $$f(\varphi_A|\varphi_B) = f_{A\cup B}\,(\varphi_{A\cup B})/f_B(\varphi_B)$$

(2.2.4) $$f_A(\varphi_A) = \int_{\times E^{|A^c|}} f \, d \otimes_n \mu.$$

Remark 2.2.1: (i) Condition (2.2.2) is what is known as the near-neighbour condition.

(ii) A Markov random field is said to be of order r $(r \geq 1)$ if the neighbours of rth order are taken into consideration.

Definition 2.2.2: We shall call $f(\varphi_t|\varphi_{\partial t})$ the local specifications of $\{X_t: t \in \Lambda\}$.

Definition 2.2.3: We say that $\{X_t: t \in \Lambda\}$ is a locally markovian field if

(i) there exists an $f > 0$ on $E \times \ldots \times E$ (n times) such that

$$P[X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n] = \int_{A_1 \times \ldots \times A_n} f \, d \otimes_n \mu \qquad \otimes_n \mu\text{-a.s.}$$

and

(2.2.5) (ii) $$f(\varphi_t|\varphi_{\{t\}^c}) = f(\varphi_t|\varphi_{\partial t}).$$

Remark 2.2.2: It is clear that a Markovian random field is also locally Markovian.

Definition 2.2.4: Let $G = (\Lambda, \Gamma)$ be a finite graph. A set of sites $K \subset \Lambda$ is called a clique if it either contains a single site or if $s,t \in K$ such that $s \in \partial t$.

Let $\mathcal{K}$ denote the family of cliques of Λ and

$$\Delta = \{\varphi_K: \varphi \in E^\Lambda \text{ and } K \in \mathcal{K}\}.$$

Example 2.2.1: In the nearest neighbour scheme of example 2.1.2 (figure 1) there are cliques of the form $\{(i,j)\}$, $\{(i-1,j), (i,j)\}$, $\{(i,j-1), (i,j)\}$ etc.... .

Definition 2.2.5: Every mapping $V: \Delta \to R$ is called a potential of Grimmett.

Remark 2.2.3: The notion of potential encountered in statistical mechanics describing the interaction between particles is much more general (see [Spitzer - 1971]). Here we consider interaction only between those particles that are near-neighbours.

Definition 2.2.6: We say that the process $\{X_t: t \in \Lambda\}$ is a Gibbs field on Λ if there exists a potential V of Grimmett such that

(2.2.6) $$P[X_{t_1} \in A_1, \dots, X_{t_n} \in A_n] = \int_{A_1 x \dots x A_n} g \, d \otimes_n \mu$$

where

(2.2.7) $$g(\varphi) = c \exp\{\sum_{K \in \mathcal{K}} V(\varphi_K)\}$$

and $A_1, \dots, A_n \in \mathcal{E}$ and $\Lambda = \{t_1, \dots, t_n\}$.

Theorem 2.2.1: For $X = \{X_t: t \in \Lambda\}$ taking values in $(E, \mathcal{E})$ the following three statements are equivalent:

(i) X is a Markovian field;

(ii) X is a locally Markovian field;

(iii) X is a Gibbs field.

Before we could give a proof of the above theorem, we need two lemmas:

Lemma 2.2.1: If f is a density of a locally Markovian field (definition 2.2.3 (i)) then $\forall\, t \in \Lambda$ $s \notin \partial t$, $C = \{s,t\}^c$, $\varphi, \varphi', \psi, \Pi, \Pi' \in E^\Lambda$ we have $\otimes_n \mu$-a.s:

(2.2.8) $$\frac{f(\varphi_t,\psi_C,\Pi_s)}{f(\varphi_t',\psi_C,\Pi_s)} = \frac{f(\varphi_t,\psi_C,\Pi_s')}{f(\varphi_t',\psi_C,\Pi_s')}$$

for an appropriate enumeration of Λ.

Proof: By definition:

$$\frac{f(\varphi_t,\psi_C,\Pi_s)}{f(\varphi_t',\psi_C,\Pi_s)} = \frac{f(\varphi_t|\psi_C,\Pi_s)}{f(\varphi_t'|\psi_C,\Pi_s)}$$

$$= \frac{f(\varphi_t|\psi_C)}{f(\varphi_t'|\psi_C)} \quad \underset{n}{\otimes}\ \mu\text{-a.s.}$$

(by Markovian property).

i.e., $\frac{f(\varphi_t,\psi_C,\Pi_s)}{f(\varphi_t',\psi_C,\Pi_s)}$ is independent of Π_s. Consequently we have

$$\frac{f(\varphi_t,\psi_C,\Pi_s)}{f(\varphi_t',\psi_C,\Pi_s)} = \frac{f(\varphi_t,\psi_C,\Pi_s')}{f(\varphi_t',\psi_C,\Pi_s')} \quad \underset{n}{\otimes}\ \mu\text{-a.s.}$$

as was desired.

Lemma 2.2.2: There exists a μ-null set $N \subset E$ such that $\forall\ \theta \in E\backslash N$, $\forall\ s,t \in \Lambda$, $s \notin \partial t$ $B \subset \Lambda\backslash\{s,t\}$ we have $\underset{n}{\otimes}\ \mu$-a.s. $\forall\ \varphi \in E^\Lambda$

(2.2.() $$\frac{f(\varphi_s,\varphi_B,\theta_\varphi^{(B\cup\{s,t\})^c},\ \varphi_t)}{f(\theta_\varphi^{\{s\}},\ \varphi_B,\ \theta_\varphi^{(B\cup\{s,t\})^c},\ \varphi_t)} = \frac{f(\varphi_s,\varphi_B,\theta_\varphi^{(B\cup\{s,t\})^c},\ \theta_\varphi^{\{t\}})}{f(\theta_\varphi^{\{s\}},\varphi_B,\theta_\varphi^{(B\cup\{s,t\})^c},\theta_\varphi^{\{t\}})}$$

Proof: By lemma 2.2.1, there exists a $\underset{n}{\otimes}\ \mu$-null set M in $E\times\ldots\times E$ (n times) such that

$$\frac{f(\varphi_s,\psi_C,\Pi_t)}{f(\varphi_s',\psi_C,\Pi_t)} = \frac{f(\varphi_s,\psi_C,\Pi_t')}{f(\varphi_s',\psi_C,\Pi_t')}$$

whenever (φ_s,ψ_C,Π_t), $(\varphi_s',\psi_C,\Pi_t)$, $(\varphi_s,\psi_C,\Pi_t')$, $(\varphi_s',\psi_C,\Pi_t')$ do not belong to M.

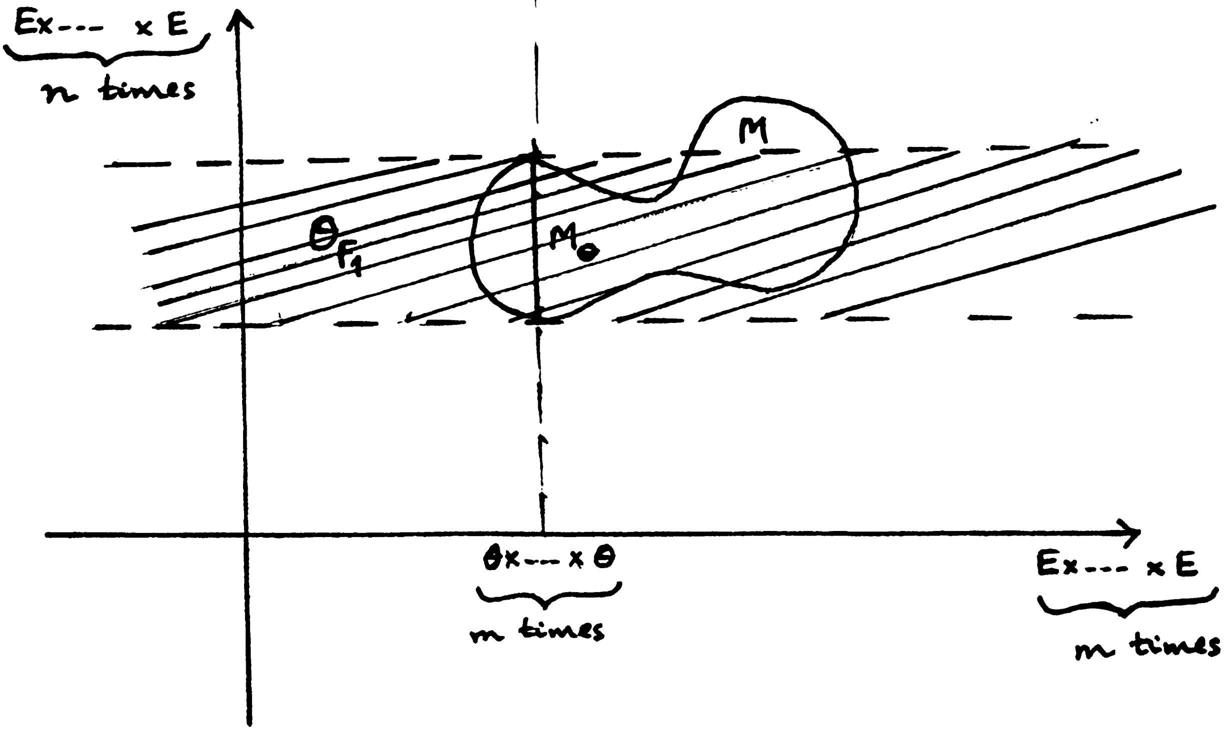
Consequently in order to justify equality (2.2.9) we need to show that each of the function in the four parentheses of the equality do not belong to M.

Let $$\theta_{F_1} = \{\varphi \in E^\Lambda \mid (\varphi_s,\varphi_B,\theta_\varphi^{(B\cup\{s,t\})^c},\varphi_t) \in M\}.$$

θ_{F_2}, θ_{F_3}, θ_{F_4} are defined similarly using the other three terms appearing in the equality (2.2.9).

Let $$m = |(B \cup \{s,t\})^c|$$

We then have the following scheme:



Since μ is σ-finite, we have

$$0 = \underset{n}{\otimes}\, \mu(M) = \int \underset{n-m}{\otimes}\, \mu(M_\theta)\, d \underset{m}{\otimes}\, \mu(\theta,\ldots,\theta)$$

implies

$$\underset{n-m}{\otimes}\, \mu(M_\theta) = 0 \qquad \mu\text{-a.s.}(\theta)$$

implies

$$\underset{n}{\otimes}\, \mu(\theta_{F_1}) = 0 \qquad \mu\text{-a.s.}(\theta)$$

Likewise we get

$$\underset{n}{\otimes}\, \mu(\theta_{F_i}) = 0 \qquad \mu\text{-a.s.}(\theta);\ i = 2,3,4$$

Therefore if $\theta \in E\setminus N$ each of the functions appearing in the four parentheses of equality 2.2.9, do not belong to M except for a null set which gives the desired conclusion of the lemma.

Proof of theorem 2.2.1:

1°: (i) $\Rightarrow$ (ii)

Follows from remark 2.2.2.

2°: (ii) $\Rightarrow$ (iii)

Given that X is locally Markovian we need to show that X is a Gibbs field.

Let $N \subset E$ be a μ-null set as in lemma 2.2.2 and let $\theta \in E\setminus N$ be fixed.

Let $U\colon\ E^\Lambda \to R$ be defined by

$$(2.2.10)\quad U(\varphi) = \sum_{B \subset \theta_{L(\varphi)}} (-1)^{|\theta_{L(\varphi)}\setminus B|} \log f(\theta_\varphi^B) \qquad \forall\, \varphi \in E^\Lambda.$$

Let now $\varphi \in E^{\Lambda}$ be fixed and $A \subseteq \theta_{L(\varphi)}$. Then

$$(2.2.11) \qquad U(\theta_{\varphi}^{A}) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \log f(\theta_{\varphi}^{B})$$

The set of subsets of $\theta_{L(\varphi)}$ is partially ordered by inclusion and its Mobius function is given by

$$\eta(B,A) = (-1)^{|A \setminus B|} \qquad \forall\ A,B \subseteq \theta_{L(\varphi)}, B \subseteq A$$

Consequently we can write (2.2.11) in the form

$$U(\theta_{\varphi}^{A}) = \sum_{B \subseteq A} \eta(B,A) \log f(\theta_{\varphi}^{B})$$

which yields

$$(2.2.12) \qquad \log f(\theta_{\varphi}^{A}) = \sum_{B \subseteq A} U(\theta_{\varphi}^{B})$$

(by Mobius inversion theorem)

<u>Claim:</u> $U(\theta_{\varphi}^{A}) = 0$ unless A is a clique.

For suppose $\phi \neq A \in K$ so that A contains two sites s and t such that $s \notin \partial t$

(2.2.11) can be rewritten as:

$$U(\theta_{\varphi}^{A}) = \sum_{\substack{B \subseteq A \\ s,t \in B}} (-1)^{|A \setminus B|} \log f(\theta_{\varphi}^{B}) + \sum_{\substack{B \subseteq A \\ s \in B, t \notin B}} (-1)^{|A \setminus B|} \log f(\theta_{\varphi}^{B})$$

$$+ \sum_{\substack{B \subseteq A \\ s \notin B, t \in B}} (-1)^{|A \setminus B|} \log f(\theta_{\varphi}^{B}) + \sum_{\substack{B \subseteq A \\ s,t \notin B}} (-1)^{|A \setminus B|} \log f(\theta_{\varphi}^{B})$$

$$= \sum_{B \subseteq A \setminus \{s,t\}} (-1)^{|A \setminus B|} [\log f(\theta_\varphi^{B \cup \{s,t\}}) - \log f(\theta_\varphi^{B \cup \{s\}})$$

$$- \log f(\theta_\varphi^{B \cup \{t\}}) + \log f(\theta_\varphi^{B})]$$

$$= \sum_{B \subseteq A \setminus \{s,t\}} (-1)^{|A \setminus B|} \log [\frac{f(\theta_\varphi^{B \cup \{s,t\}})}{f(\theta_\varphi^{B \cup \{t\}})} / \frac{f(\theta_\varphi^{B \cup \{s\}})}{f(\theta_\varphi^{B})}]$$

$$= 0 \quad \underset{n}{\otimes} \mu\text{-a.s. (by lemma 2.2.2)}$$

Thus (2.2.12) gives in particular for $A = \theta_{L(\varphi)}$

$$f(\varphi) = \exp\{ \sum_{\substack{K \subseteq \theta_{L(\varphi)} \\ K \in \mathcal{K}}} U(\theta_\varphi^K) + U(\theta_\varphi \phi)\} \underset{n}{\otimes} \mu\text{-a.s.}$$

Hence:

$$f(\varphi) = \text{const} \ \exp\{ \sum_{\substack{k \subseteq \theta_{L(\varphi)} \\ K \in \mathcal{K}}} U(\theta_\varphi^K)\} \underset{n}{\otimes} \mu\text{-a.s.}$$

which can be written as

$$f(\varphi) = \text{const} \ \exp\{ \sum_{K \in \mathcal{K}} V(\varphi_K)\} \underset{n}{\otimes} \mu\text{-a.s.}$$

where

$$V(\varphi_K) = \begin{cases} U(\theta_\varphi^K) & \text{if} \quad \theta_{L(\theta_\varphi^K)} = K \\ 0 & \text{otherwise.} \end{cases}$$

Thus, there exists a potential function $V: \Delta \to R$ such that

$$P[X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n] = \int_{A_1 \times \ldots \times A_n} f \, d \underset{n}{\otimes} \mu$$

with

$$f(\varphi) = \text{const} \ \exp\{ \sum_{K \in \mathcal{K}} V(\varphi_K) \} \qquad \underset{n}{\otimes} \mu\text{-a.s.}$$

implies X is a Gibbs field.

3°: (iii) $\Rightarrow$ (i)

Only need to show that condition (2.2.2) of definition (2.2.1) is satisfied.

i.e., need to show that $\forall \ A \subseteq \Lambda, \ \varphi \in E^{\Lambda}$

$$g(\varphi_A | \varphi_{A^c}) = g(\varphi_A | \varphi_{\partial A}) \quad \underset{n}{\otimes} \mu\text{-a.s.}$$

For simplicity we shall omit $\underset{n}{\otimes} \mu$-a.s. Using (2.2.3) and (2.2.4) we have

$$g(\varphi_A | \varphi_{A^c}) = \frac{g(\varphi)}{\int_{\underset{|A|}{\times} E} g(\varphi) \, d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

$$= \frac{\cancel{c} \exp\{ \sum_{\substack{K \cap A \neq \phi \\ K \in \mathcal{K}}} V(\varphi_K) \} \cancel{\exp\{ \sum_{\substack{K \cap A = \phi \\ K \in \mathcal{K}}} V(\varphi_K) \}}}{\cancel{c} \int_{\underset{|A|}{\times} E} \exp\{ \sum_{\substack{K \cap A \neq \phi \\ K \in \mathcal{K}}} V(\varphi_K) \} \cancel{\exp\{ \sum_{\substack{K \cap A = \phi \\ K \in \mathcal{K}}} V(\varphi_K) \}} d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

$$= \frac{\exp\{ \sum_{\substack{K \cap A \neq \phi \\ K \in \mathcal{K}}} V(\varphi_K) \}}{\int_{\underset{|A|}{\times} E} \exp\{ \sum_{\substack{K \cap A \neq \phi \\ K \in \mathcal{K}}} V(\varphi_K) \} d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

Similarly we get:

$$g(\varphi_A|\varphi_{\partial A}) = \frac{\int\limits_{\substack{xE\\|\overline{A}^c|}} g(\varphi)\,d \underset{|\overline{A}^c|}{\otimes} \mu(\varphi_{\overline{A}^c})}{\int\limits_{\substack{xE\\|A|}} \{\int\limits_{\substack{xE\\|\overline{A}^c|}} g(\varphi)\,d \underset{|\overline{A}^c|}{\otimes} \mu(\varphi_{\overline{A}^c})\}\, d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

(Here $\overline{A} = A \cup \partial A$)

$$= \frac{c \int\limits_{\substack{xE\\|\overline{A}^c|}} \exp\{\sum\limits_{\substack{K\cap A\neq\phi\\K\in\mathcal{K}}} V(\varphi_K)\} \exp\{\sum\limits_{\substack{K\cap A=\phi\\K\in\mathcal{K}}} V(\varphi_K)\}\, d \underset{|\overline{A}^c|}{\otimes} \mu(\varphi_{\overline{A}^c})}{c \int\limits_{\substack{xE\\|A|}} \{\int\limits_{\substack{xE\\|\overline{A}^c|}} \exp\{\sum\limits_{\substack{K\cap A\neq\phi\\K\in\mathcal{K}}} V(\varphi_K)\} \exp\{\sum\limits_{\substack{K\cap A=\phi\\K\in\mathcal{K}}} V(\varphi_K)\}\, d \underset{|\overline{A}^c|}{\otimes} \mu(\varphi_{\overline{A}^c})\} d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

$$= \frac{\exp\{\sum\limits_{\substack{K\cap A\neq\phi\\K\in\mathcal{K}}} V(\varphi_K)\}}{\int\limits_{\substack{xE\\|A|}} \exp\{\sum\limits_{\substack{K\cap A\neq\phi\\K\in\mathcal{K}}} V(\varphi_K)\}\, d \underset{|A|}{\otimes} \mu(\varphi_A)}$$

(since the other term does not depend upon φ_A)

Therefore it follows that

$$g(\varphi_A|\varphi_{A^c}) = g(\varphi_A|\varphi_{\partial A}) \quad \underset{n}{\otimes} \mu\text{-a.s.}$$

1°, 2° and 3° together imply that the three statements of the theorem are equivalent.

Remark 2.2.4: The above theorem shows that the set of Markov (or locally Markov) fields and Gibbs states with nearest neighbour potentials are the same for any finite graph.

Remarks 2.2.5: (i) Grimmett (1973) has proved the part (ii) $\Rightarrow$ (iii) of theorem 2.2.1 in the case where E is countable. One needs lemma 2.2.2 to pass to the general case of carnal (1979).

(ii) In the case when E is discrete, Besag (1974) gives a simple alternative proof of the part (ii) ⇒ (iii). See theorem 2.2.2 of this section.

(iii) For $E = \{0,1\}$ Preston (1974) gives a complete discussion. For details, one is referred to his paper.

<u>Corollary to theorem 2.2.1</u>: If X is locally Markovian then it is completely specified by its local specifications.

<u>Proof</u>: It was seen in theorem 2.2.1 that X is completely described by the U-functions defined by (2.2.10). Therefore, it suffices to show that the local specifications determine the $U(\theta_\varphi^K)$ $\forall$ clique $K \in \mathcal{K}$.

Let us consider θ and $\varphi \equiv \theta$ on Λ to be identical. Since X is locally Markovian:

$$f(\theta_t|\theta_{\partial t}) = f(\theta_t|\theta_{\{t\}^c})$$

$$= \frac{f(\theta)}{\int_E f(\theta_\varphi^t)\, d\,\mu(\varphi_t)}$$

$$= \frac{\text{const}}{\int_E c \exp U(\theta_\varphi^t)\, d\,\mu(\varphi_t)}$$

Likewise it can be seen that

$$f(\varphi_t|\theta_{\partial t}) = \frac{c \exp U(\theta_\varphi^t)}{\int_E c \exp U(\theta_\varphi^t)\, d\,\mu(\varphi_t)}$$

Consequently it follows that

$$U(\theta_\varphi^t) = \log \frac{f(\varphi_t|\theta_{\partial t})}{f(\theta_t|\theta_{\partial t})}$$

Therefore we know the form of $U(\theta_\varphi^K)$ $\forall$ clique K with a single element, $K = \{t\}$. The proof is then completed by induction and one is referred to Carnal (1979) for completion of this proof.

Remark 2.2.6: Carnal (1979) has also discussed the case when Λ is a finite subset of a circle.

2.2.1 A special case of theorem 2.2.1:

Let there be n sites in Λ labelled $1,2,\ldots,n$ and $\{X_i: i = 1,2,\ldots,n\}$ be a family of random variables associated with these n sites taking values in a discrete space E. Let $X = \{X_i: i = 1,2,\ldots,n\}$ be a Markov field. Then the near-neighbour condition (2.2.2) of definition (2.2.1) implies that $P[X_i|\text{all other site values}] = P[X_i|\text{neighbours of site } i]$.

Besag (1974) gives a representation of the probability structure $P(\vec{X})$ of $\vec{X} = (x_1,\ldots,x_n)$ through near-neighbour system in the simplest form. Following his paper, let us assume that if $P(x_i) > 0$ for each i then $P(\vec{x}) > 0$ (known as the positivity condition).

Let $\Omega^* = \{\vec{X} : P(\vec{X}) > 0\}$ be the sample space of all realizations of the system. In what follows it will prove convenient to consider the representation for the ratio $P(\vec{X})/P(\vec{0})$.

Define $V(\vec{X}) = \ell n\{P(\vec{X})/P(\vec{0})\}$.

Lemma 2.2.3: There exists an expansion of $V(\vec{X})$ unique on Ω^* given by

$$V(\vec{X}) = \sum_{1\leq i\leq n} x_i F_i(x_i) + \sum_{1\leq i<j\leq n} x_i x_j F_{i,j}(x_i,x_j)$$

(2.2.13)

$$+\ldots+ x_1 x_2 \ldots x_n\, F_{1,2,\ldots,n}(x_1,\ldots,x_n)$$

Proof: The F-functions are determined inductively. For example:

$$x_i F_i(x_i) = V(0,\ldots,0,x_i,0,\ldots,0) - V(\vec{0})$$

with analogous difference formulae for higher order F-functions.

Theorem 2.2.2: [Besag - 1974].

$V(\vec{X})$ defined by (2.2.13) gives a valid probability structure to the Markovian random field taking values in a discrete space E provided the functions $F_{i,j,\ldots,s}(x_i,x_j,\ldots,x_s)$ are non-zero which holds if and only if the sites $i,j,\ldots,s$ form a clique. Subject to this restriction, the F-functions may be chosen arbitarily.

Remark 2.2.7: The above theorem says that given the neighbours of each site, we can readily write down the most general form of $V(\vec{X})$ and consequently of $P(\vec{X})$, the probability structure of the Markovian random field taking its values in a discrete space.

2.3. Characterization of spatial schemes:

In practice, we shall often find that the points or sites occur in a finite region of the euclidean space and often fall into two categories: those which are internal to the system and those which form their boundary (or boundaries). In the analysis of spatial patterns one is interested in the behaviour of the system outside the region of their occurrence. The problems at the boundary may be handled by considering the joint distribution of the internal site variables conditional upon fixed observed boundary values. Thus, in constructing spatial schemes to study the behaviour of the system we need only to specify the neighbours of a given set of sites and the associated conditional probability structure for each of the sites. This is exactly what is specified by Markovian random fields. Consequently, in order to study the behaviour of a spatial pattern outside the region of occurrence it is enough to consider schemes that represent Markov random fields. Thus, in what follows we shall be considering only those spatial schemes that satisfy Markovian property and we shall use the notations of section 2.2.

Proposition 2.3.1: For a Markov random field with no interaction the representing σ-finite measure μ is the Poisson measure.

Proof: By remark 2.2.4 and theorem 2.2.1 the density of a Markov random field with respect to any σ-finite measure μ is given by

$$P[X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n] = c \int_{A_1 x \ldots x A_n} \exp\{\sum_K V(\varphi_K)\} d \underset{n}{\otimes} \mu$$

where c is a constant, V is the near-neighbour potential and $A_1, \ldots, A_n \in E$ The term $\sum_K V(\varphi_K)$ is the contribution from the inter-action between the particles. Consequently, the density of a Markov random field with no interaction with respect to μ is:

$$P[X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n] = c \int_{A_1 x \ldots x A_n} d \otimes_n \mu$$

i.e.,

$$P[X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n] = c \prod_{i=1}^{n} \mu(A_i) \tag{2.3.1}$$

By proposition 1.5.1, ℓ.h.s. represents the Poisson law so that it follows that the finite dimensional distributions of a spatial scheme with no interaction has a finite dimensional Poisson law. Consequently by theorem 1.4.2 it follows that for a Markov random field with no interaction the representing measure μ is the Poisson measure.

Remark 2.3.1: Theorem 2.2.1 says that Markov random fields have an exponential density indexed by V, the potential of Grimmett with respect to any σ-finite measure μ and proposition 2.3.1 says that the representing measure μ is the Poisson measure for a Markov random field with no interaction. Consequently, it follows that in testing for randomness namely $V \equiv 0$ against the alternatives $V > 0$, the spatial scheme will have an exponential density indexed by V with respect to the Poisson measure.

CHAPTER III

INFERENCES ON RANDOMNESS

In the analysis of spatial patterns one of the problems of main interest is to determine whether or not the spatial patterns exhibit any randomness i.e. whether or not the observed pattern is a realization of a Poisson point process. This may be done through

(i) estimation of λ, the intensity of the spatial process;

and

(ii) testing hypotheses concerning the parameters namely V, Grimmett's potential that describe the spatial interaction.

For the testing hypothesis problem (section 3.2) we consider μ of Chapter 2 to be a 0-1 variable and $V(\vec{X})$ of section 2.2 to be of a particular form (N-N interaction) and compute the asymptotic distribution of log-likelihood ratio under the null hypothesis of no spatial interaction. This justifies the recent work of Besag (1974) and it has been remarked that Besag's coding method of estimation is not necessary in establishing the chi-square behaviour of the log-likelihood.

<u>3.1 Methods of estimation of λ, the expected number of individuals per unit area</u>.

In the analysis of spatial patterns, estimation of λ plays an important role because it contributes to the understanding of certain aspects of the pattern or the arrangement of individuals in space. In this section we discuss some of the methods of estimation of λ and the asymptotic properties of these estimators. The estimators are constructed so that they are at least asymptotically unbiased. There are mainly two techniques described respectively as quadrat method and distance method.

3.1.1 Quadrat Method: This method is based on field sampling and involves choosing m disjoint quadrats each of area D from the region of study.

Let $Z_i (i = 1,\ldots,m)$ denote the number of individuals in the ith quadrat. Assuming that the pattern is random by proposition 1.5.1 each $Z_i (i = 1,2,\ldots,m)$ has a Poisson distribution with parameter λD and are independent.

Consequently the sample likelihood functions based on these m quadrat counts is given by

$$(3.1.1) \qquad L_q = \frac{e^{-\lambda m D} (\lambda D)^{\sum_{i=1}^{m} Z_i}}{\prod_{i=1}^{m} Z_i!}$$

Clearly $\sum_{i=1}^{m} Z_i$ is a complete sufficient statistic.

Using (3.1.1) the maximum likelihood estimator of λ based on quadrat counts is given by

$$(3.1.2) \qquad \hat{\lambda}_{q,m} = \frac{\sum_{i=1}^{m} Z_i}{mD}$$

also known as the quadrat estimator of λ.

Remarks 3.1.1: (i) It is clear that $\hat{\lambda}_{q,m}$ given by (3.1.2) is an unbiased estimator of λ.

(ii) If the D's are not equal then Z_i's will be independent and distributed as a Poisson random variable with parameter λD_i $(i = 1,2,\ldots,m)$. In this case the quadrat estimator has the form

$$\hat{\lambda}_{q,m} = \frac{\sum_{i=1}^{m} Z_i}{\sum_{i=1}^{m} D_i}$$

which again is an unbiased estimator of λ.

Theorem 3.1.1: Assuming we have a random pattern (i.e., the pattern is a realization of a homogeneous Poisson point process) the quadrat estimator $\hat{\lambda}_{q,m}$ defined by (3.1.2) is

a) strongly consistent

i.e.,

(3.1.3) $$\hat{\lambda}_{q,m} \xrightarrow{a.s.} \lambda \quad \text{as} \quad m \to \infty.$$

b) asymptotically normally distributed: In fact

(3.1.4) $$\sqrt{m}(\hat{\lambda}_{q,m} - \lambda) \xrightarrow{\mathcal{D}} N(0,\lambda/D)$$

Proof: a) By (3.1.2)

$$\hat{\lambda}_{q,m} = \sum_{i=1}^{m} Z_i/mD$$

By proposition 1.5.1 $\{Z_i:\ i = 1,\ldots,m\}$ is a sequence of independent identically distributed (i.i.d. for short) Poisson random variables with parameter λD. Thus, using the classical theory of i.i.d. random variables it follows that

$$\hat{\lambda}_{q,m} \xrightarrow{a.s.} \lambda \quad \text{as} \quad m \to \infty.$$

b) Now:

$$\sqrt{m}(\hat{\lambda}_{q,m} - \lambda) = \sqrt{m}\ \{ \frac{\sum_{i=1}^{m} Z_i}{mD} - \lambda\}$$

$$= \frac{1}{D}\ \{\frac{1}{\sqrt{m}} \sum_{i=1}^{m} (Z_i - \lambda D)\}$$

Here $\{Z_i - \lambda D:\ i = 1,\ldots,m\}$ is a sequence of i.i.d. random variables with mean zero and variance λD. Thus, by central limit theorem it follows that $\frac{1}{\sqrt{m}} \sum_{i=1}^{m} (Z_i - \lambda D)$ is asymptotically normally distributed with mean zero and variance λD.

Hence

$$\sqrt{m}(\hat{\lambda}_{q,m} - \lambda) \xrightarrow{\mathcal{D}} N(0,\lambda/D).$$

Remark 3.1.2: (3.1.4) implies that the asymptotic variance of $\hat{\lambda}_{q,m}$ is $\frac{\lambda}{mD}$ which implies that if mD can be taken arbitrarily large then λ can be determined with high accuracy using the quadrat estimator.

3.1.3 Distance Method

Distance method involves estimation and testing of parameters for a spatial process based on some kind of distance measurements. Various distance measures have been studied in the literature. For example, the distance measured may be from an arbitrarily chosen (explained below) point to its nearest neighbour, second nearest neighbour ,...,rth nearest neighbour etc.; the distance measured using T-square sampling introduced by Besag and Gleaves (1973) and many more. For more details on T-square sampling the interested reader is referred to the paper by Diggle, Besag and Gleaves (1976).

In this study we will confine our attention only to nearest neighbour (N-N for short) distances which probably is one of the simplest distance measures.

Before we can give the estimator based on N-N distance measurements, we need the following proposition concerning the distribution of N-N distances.

Proposition 3.1.1: Suppose the points come from a homogeneous Poisson point process with intensity λ. Let X denote the distance of an arbitrarily chosen point from its N-N. Then the transformed variable $Y = V(X)$ where $V(X)$ denotes the volume of a sphere centered at the chosen point and with radius X, has an exponential distribution with parameter λ.

Proof: Choose an arbitrary point from a realization of a homogeneous Poisson point process with intensity λ.

For a Poisson process, the probability of capturing exactly k individuals in a sphere of radius x is given by

$$\frac{e^{-\lambda V(x)}(\lambda\ V(x))^k}{k!} \quad k = 0,1,2,\ldots$$

Now the event $[X > x]$ denotes the event that no point is captured in the sphere of radius x and centered at the chosen point.

Thus,

$$P[X > x] = e^{-\lambda\, V(x)} \qquad x > 0.$$

This implies

$$F_X(x) = P[X \leq x] = 1 - e^{-\lambda\, V(x)} \qquad x > 0$$

implies $dF_X(x) = \lambda\, e^{-\lambda\, V(x)}\, V'(x)dx \quad x > 0$. Consequently, the transformed variable $Y = V(X)$ has the distribution function given by

$$dG_Y(y) = \lambda \exp(-\lambda y)dy \qquad y > 0$$

i.e., Y has the probability density function

$$g(y) = \lambda \exp(-\lambda y) \quad y > 0$$

which is an exponential density with parameter λ.

<u>Remark 3.1.3</u>: If in proposition 3.1.1, X denotes the distance from an arbitrarily chosen point to its rth $(r \geq 1)$ nearest neighbour, then the event $[X > x]$ describes that there are at most $(r-1)$ points captured in the sphere of radius x, centered at the chosen point so that

$$P[X > x] = \sum_{k=0}^{r-1} \frac{e^{-\lambda\, V(x)}\, [\lambda V(x)]^k}{k!}$$

implies

$$F_X(x) = 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda\, V(x)} (\lambda\, V(x))^k}{k!}$$

Thus, X has the p.d.f. given by

$$f(x) = \frac{\lambda^r \exp(-\lambda\ V(x))(V(x))^{r-1} V'(x)}{(r-1)!} \quad x > 0$$

Consequently, the transformed variable $Y = V(X)$ has the p.d.f. given by

$$g(y) = \frac{\lambda^r\ e^{-\lambda y}\ y^{r-1}}{\Gamma(r)} \quad y > 0$$

which is a gamma density with parameters r and λ.

Near-neighbour estimator of λ based on distance measurements:

Sampling scheme for the distance approach involves choosing arbitrarily a set of n sample points from the region of study. By proposition 1.5.1, the points are from a realization of a homogeneous Poisson point process with intensity λ.

Let X_j denote the distance of the jth point from its nearest neighbour. By proposition 3.1.1, it then follows that the transformed variable $Y_j = V(X_j)$, $V(X_j)$ being the volume of the sphere centered at the chosen point and with radius X_j, has an exponential density with parameter λ. The points are so chosen that the Y_j's are independent. The sample likelihood function based on n N-N distance measurements is:

$$L_d = \lambda^n \exp(-\lambda \sum_{j=1}^{n} Y_j) \tag{3.1.6}$$

Clearly $\sum_j Y_j$ is a complete sufficient statistic and the maximum likelihood estimator of λ using (3.1.6) based on distance measurements is given by

(3.1.7) $$\hat{\lambda}_{d,n} = n \Big/ \sum_{j=1}^{n} Y_j$$

We shall refer to $\hat{\lambda}_{d,n}$ as the N-N estimator of λ.

Remark 3.1.4: In the preceding formulation for $\hat{\lambda}_{d,n}$ one may also consider point-to-plant or plant-to-plant distances. However, as remarked by Pielou (1959) a measure based on plant-to-plant distances may not reveal any nonrandomness in the spatial distribution at all since plants may often be present in clumps or clumps of clumps etc...

Theorem 3.1.2: Assuming we have a random pattern (i.e. pattern is a realization of a homogeneous Poisson point process) the N-N estimator $\hat{\lambda}_{d,n}$ given by (3.1.7) is:

a) strongly consistent

i.e.,

(3.1.8) $$\hat{\lambda}_{d,n} \xrightarrow{a.s.} \lambda \quad \text{as} \quad n \to \infty.$$

b) asymptotically unbiased.

In fact:

(3.1.9) $$E(\hat{\lambda}_{d,n}) = \frac{n\lambda}{n-1} \text{ , and}$$

(3.1.10) $$Var(\hat{\lambda}_{d,n}) = \frac{n^2\lambda^2}{(n-1)^2(n-2)}$$

Proof: By proposition 3.1.1 $\{Y_j: \; j = 1,\ldots,n\}$ is a sequence of i.i.d. exponential random variables with parameter λ.

Therefore by strong law of large numbers we have:

$$\frac{\sum_{j=1}^{m} Y_j}{n} \xrightarrow{a.s.} E(Y_1) = \frac{1}{\lambda} \quad \text{as} \quad n \to \infty$$

and consequently it follows that

$$\hat{\lambda}_{d,n} \xrightarrow{a.s.} \lambda \quad \text{as} \quad n \to \infty.$$

b) Let now $Y = \sum_{j=1}^{n} Y_j$ so that $\hat{\lambda}_{d,n} = nY^{-1}$ Since each $Y_j (j = 1,\ldots,n)$ is exponential with parameter λ it follows that Y is a Gamma random variable with parameters n and λ.

i.e. $$P[Y \leq t] = \int_0^t \frac{\lambda^n y^{n-1} e^{-\lambda y}}{\Gamma(n)}\, dy$$

Consequently it follows that

$$F_{Y^{-1}}(t) = P[Y^{-1} \leq t] = 1 - \int_0^{t^{-1}} \frac{\lambda^n e^{-\lambda y} y^{n-1}}{\Gamma(n)}\, dy$$

For $r > 0$, it can easily be seen that

$$E(Y^{-r}) = r \int_0^{\infty} t^{r-1}[1 - F_{Y^{-1}}(t)]dt$$

implies: $$E(Y^{-1}) = \lambda/(n-1)$$

and $$E(Y^{-2}) = \lambda^2/(n-1)(n-2)$$

Hence it follows that

$$E(\hat{\lambda}_{d,n}) = n\lambda/(n-1) \quad \text{and}$$

$$\text{Var}(\hat{\lambda}_{d,n}) = \frac{n^2\lambda^2}{(n-1)^2(n-2)}$$

clearly (3.1.9) implies the asymptotic unbiasedness of $\hat{\lambda}_{d,n}$.

Remarks 3.1.5: (i) (3.1.9) says that the N-N estimator $\hat{\lambda}_{d,n}$ is slightly biased. However if the spatial distribution is not random (i.e. if it is not a realization of a homogeneous Poisson point process) this estimator may give serious bias.

(ii) Using the classical theory on i.i.d. random variables it follows that

$$\sqrt{n}\left(\frac{\sum_j Y_j}{n} - \frac{1}{\lambda}\right) \xrightarrow{\mathcal{D}} N(0,1/\lambda^2)$$

Consequently it follows that

$$\sqrt{n}(\hat{\lambda}_{d,n} - \lambda) \xrightarrow{\mathcal{D}} N(0,\lambda^2)$$

(iii) Using the quadrat estimator $\hat{\lambda}_{q,m}$ and the N-N estimator $\hat{\lambda}_{d,n}$ let us define an index:

$$\alpha = \frac{\hat{\lambda}_{q,m}}{\hat{\lambda}_{d,n}} = \frac{\sum_{i=1}^{m} Z_i/mD}{n/\sum_{j=1}^{n} Y_j}$$

Under randomness assumption, it was noted earlier that both $\hat{\lambda}_{q,m}$ and $\hat{\lambda}_{d,n}$ converge to λ so that $\alpha \to 1$. Thus if one calculates α from the observed data and finds that it differs significantly from 1 then it can be assumed that the spatial distribution is not random. Also in an aggregated population we would expect higher values to α and in a regularly dispersed population we would expect low values to α. Pielou (1959) has given approximate confidence intervals for α. For more details one is referred to Pielou's paper.

(iv) Another way to test for randomness: In the formulation for N-N estimator $\hat{\lambda}_{d,n}$ it was noted that $Y = \sum_{j=1}^{n} Y_j$ has a Gamma distribution with parameters n and λ under the randomness assumption. Consequently under randomness $\gamma = 2\lambda \sum Y_j$ has a chi-square distribution with $2n$ degrees of freedom. This γ may be used to test for randomness rather than α.

3.1.3 Estimator of λ based on both quadrat counts and N-N distance measurements:

By considering two independent realizations of the Poisson point process with intensity λ, one can have two independent sets of data namely m quadrat counts and n N-N distance measurements. Thus it seems reasonable to use these two independent sets of data to look at an estimator of λ and the sample likelihood function based on m quadrat counts and n distance measurements is therefore:

$$L_{q,d} = \lambda^{n+\sum_{i=1}^{m} Z_i} \exp\{-\lambda(mD + \sum_j Y_j)\} \frac{D^{\sum_i Z_i}}{\prod_i Z_i!} \tag{3.1.11}$$

where Z_i $(i = 1,\ldots,m)$ are the quadrat counts and $Y_j = V(X_j)$ representing the volume of the sphere with radius X_j - the distance from the jth point to its N-N, are the distance measurements.

In (3.1.11) the r.h.s. is a product of two exponential families of distributions and hence is itself an exponential family of distributions. Further, it can be noted from (3.1.11) that there is no single sufficient statistic for λ but $(\sum Z_i, \sum Y_j)$ is a sufficient statistic pair. Using (3.1.11) the maximum likelihood estimator of λ based on both quadrat counts and distance measurements is:

(3.1.12) $$\hat{\lambda}_{m,n} = \frac{n + \sum_i Z_i}{mD + \sum Y_j}$$

<u>Theorem 3.1.3</u>: Assuming we have a random pattern (in the sense that it is a realization of a homogeneous Poisson point process) and $n/mD \to \infty$ as $n,m \to \infty$ the estimator $\hat{\lambda}_{m,n}$ defined by (3.1.12) is:

a) strongly consistent;

b) asymptotically normally distributed.

In fact:

(3.1.13) $$\sqrt{n}\,(\hat{\lambda}_{m,n} - \lambda) \overset{\mathcal{D}}{\to} N(0,\lambda^2/2)$$

<u>Proof</u>: a) the hypothesis that $\frac{n}{mD} \to \lambda$ as $n,m \to \infty$ is redundant for this part.

By (3.1.12): $$\hat{\lambda}_{m,n} = \frac{n + \sum_i Z_i}{mD + \sum_j Y_j}$$

$$= \frac{\frac{n}{mD} + \frac{1}{mD} \sum_{i=1}^{m} Z_i}{1 + \frac{n}{mD}\frac{1}{n} \sum_{j=1}^{n} Y_j}$$

By proposition 1.5.1 $\{Z_i\colon\ i = 1,2,\ldots,m\}$ is a sequence of i.i.d. Poisson random variables with parameter λD so that by SLLN:

(3.1.14) $$\frac{1}{mD} \sum_{i=1}^{m} Z_i \xrightarrow{a.s.} \lambda \quad \text{as} \quad m \to \infty$$

Similary by proposition 3.1.1 $\{Y_j\colon\ j = 1,\ldots,n\}$ is a sequence of i.i.d. exponential random variables with parameter λ so that again by SLLN:

(3.1.15) $$\frac{1}{n}\sum_{j=1}^{n} Y_j \xrightarrow{a.s.} \frac{1}{\lambda} \quad \text{as} \quad n \to \infty$$

Thus (3.1.14) and (3.1.15) imply that as $n,m \to \infty$

$$\hat{\lambda}_{m,n} \xrightarrow{a.s.} \lambda$$

b) $$\sqrt{n}\,(\hat{\lambda}_{m,n} - \lambda) = \sqrt{n}\,\{\frac{n + \sum Z_i}{mD + \sum Y_i} - \lambda\}$$

Rearranging terms this can be rewritten as:

$$\sqrt{n}\,(\hat{\lambda}_{m,n} - \lambda) = \frac{n}{mD}\{\frac{1}{\sqrt{n}}\sum_{i=1}^{m}(Z_i - \lambda D) - \frac{\lambda}{\sqrt{n}}\sum_{1}^{n}(Y_j - \frac{1}{\lambda})\}$$

$$= \frac{\frac{n}{mD}\{\sqrt{\frac{m}{n}}\cdot\frac{1}{\sqrt{m}}\sum_{i=1}^{m}(Z_i - \lambda D) - \frac{\lambda}{\sqrt{n}}\sum(Y_j - \frac{1}{\lambda})\}}{1 + \frac{n}{mD}\,\frac{1}{n}\sum_j Y_j}$$

Now (i) $\{Z_i - \lambda D: \ i = 1,\ldots,m\}$ is a sequence of i.i.d. random variables wiht mean zero and variance λD;

(ii) $\{Y_j - \frac{1}{\lambda}: \ j = 1,\ldots,n\}$ is a sequence of i.i.d. random variables with mean zero and variance $\frac{1}{\lambda^2}$.

Thus, by multivariate central limit theorem it follows that

$$(\frac{1}{\sqrt{m}}\sum(Z_i - \lambda D), \frac{1}{\sqrt{n}}\sum(Y_j - \frac{1}{\lambda}))$$

$$\xrightarrow{\mathcal{D}} N(\vec{0}, \begin{pmatrix} \lambda D & 0 \\ 0 & 1/\lambda^2 \end{pmatrix})$$

using the independence of Z_i's, Y_j's and $\frac{n}{mD} \to \lambda$ as $n,m \to \infty$ we have

$$\sqrt{n}(\hat{\lambda}_{m,n} - \lambda) \xrightarrow{\mathcal{D}} N(0,\lambda^2/2)$$

Remark 3.1.7: (3.1.13) implies that the asymptotic variance of $\hat{\lambda}_{m,n}$ is $\lambda^2/2n$ which is less than the asymptotic variance λ/mD of $\hat{\lambda}_{q,m}$ and also than the asymptotic variance λ^2/n of the unbiased version of the N-N estimator. In practice however either $\hat{\lambda}_{q,m}$ or $\hat{\lambda}_{m,n}$ is used and of course the choice between $\hat{\lambda}_{q,m}$ and $\hat{\lambda}_{m,n}$ will affect the performance of the test.

In the following section we consider a particular subclass of spatial Markov random fields and become more specific with some of the spatial schemes generated by this subclass. In the later part of the section we look at a test for randomness for spatial binary schemes in this subclass.

3.2 Spatial schemes generated by a subclass of Markov random fields and testing of hypothesis for binary models:

3.2.1 One dimensional problem:

Let there be n sites labelled $1,2,\ldots,n$ and a set of neighbours for each site. Let X_i: $i = 1,2,\ldots,n$ denote the site variables. Then following section 2.2.1 a class of valid probability structure associated with these site variables is given by

$$(3.2.1) \qquad P(\vec{X}) = P(\vec{0}) \exp V(\vec{X})$$

where

$$V(\vec{X}) = \sum_{1\le i\le n} x_i\, F_i(x_i) + \sum_{1\le i<j\le n} x_i x_j\, F_{i,j}(x_i,x_j)$$

$$+ \ldots + x_1 x_2 \ldots x_n\, F_{1,2,\ldots,n}(x_1,\ldots,x_n)$$

with $F_{i,j,\ldots,s}(x_i,x_j,\ldots,x_s)$ non-zero if and only if the sites $i,j,\ldots,s$ form a clique. Subject to this restriction, the F-functions may be chosen arbitrarily.

We shall use the function $p_i(\cdot)$ to denote the conditional probability distribution (or density function) of X_i given all other site values. However, by the Markovian property $p_i(\cdot)$ will be a function of x_i and of the values at sites neighbouring site i. Within this framework we consider a particular subclass of Markov random fields for which $V(\vec{X})$ is well defined and has the form:

$$V(\vec{X}) = \sum_i x_i F_i(x_i) + \sum_{i,j} \beta_{i,j} x_i x_j \tag{3.2.2}$$

where $\beta_{i,j} = 0$ unless sites i and j are neighbours of each other. i.e., in particular the only non-zero parameters are those associated with the cliques consisting of single sites and of pairs of sites.

Spatial Markov random fields whose probability structure is given by (3.2.1) with $V(\vec{X})$ given by (3.2.2) are known as <u>Auto-models</u> and $\{\beta_{i,j}\}$ are called the parameters of the models that describe spatial interaction between the sites.

We shall specifically be dealing with the subclass of auto-models for which $\beta_{i,j} = \beta \;\forall\; i$ and j so that β describes the spatial interaction between near-neighbour sites. In such cases, the auto-models are said to be <u>homogeneous</u>.

<u>Remark 3.2.1</u>: In view of (3.2.2), the homogeneous auto-models have conditional probability structure satisfying:

$$\frac{p_i(x_i;\ldots)}{p_i(0;\ldots)} = \exp\{x_i[F_i(x_i) + \beta \sum_j x_j]\} \tag{3.2.3}$$

where $\sum_j x_j$ will denote the sum of the values at sites neighbouring site i. The models can further be classified into auto-normal, auto-logistic, auto-binomial according as $p_i(\cdot)$ taking normal, logistic or binomial form.

3.2.2 One-dimensional auto-logistic model for binary data:

In this special case the site variables X_i take 0-1 values. For any finite system of binary variables, the only situation in which the non-zero F-functions can contribute to $V(\vec{X})$ (given by 3.2.2) are those upon which each of the arguments is unity. We may therefore replace the non-zero functions by arbitrary parameters. However, since β is the one that describes the spatial interaction, without any loss of generality we may replace $F_i(x_i)$ by a constant namely α. Thus, the spatial binary scheme has a probability structure given by (3.2.1) for which $V(\vec{X})$ has the form:

$$V(\vec{X}) = \alpha \sum_j x_i + \beta \sum_{1 \leq i \leq j \leq n} x_i x_j$$

Consequently it follows from (3.2.3) that:

$$(3.2.4) \qquad p_i(x_i;\ldots) = \frac{\exp\{x_i[\alpha + \beta \sum_j x_j]\}}{1 + \exp[\alpha + \beta \sum_j x_j]}$$

where $\sum_j x_j = x_{i-1} + x_{i+1}$ is the sum of the values at sites neighbouring site i.

The model specified by (3.2.4) is the classical logistic model and thus in this case the spatial model is known as the auto-logistic model.

Remark 3.2.2: Auto-logistic models are quite useful in practice for example in an ecological context the variables may correspond to an array of plants each of which is either infected (1) or healthy (0), or to the presence (1) or absence (0) of a plant at a site. Moreover, the models having once been established are easy to interpret.

Remark 3.2.3: A homogeneous first-order (or nearest neighbour) scheme for zero-one variables for a rectangular lattice with sites labelled by integer pairs (i,j) is given by [Besag - 1974]:

$$P[x_{ij} | x_{i-1,j}, x_{i+1,j}, x_{i,j-1}, x_{i,j+1}]$$

$$= \frac{\exp[x_{ij}\, t^*_{i,j}]}{1 + \exp[t^*_{i,j}]} \tag{3.2.5}$$

where $t^*_{i,j} = \alpha + \beta_1(x_{i-1,j} + x_{i+1,j}) + \beta_2\,(x_{i,j-1} + x_{i,j+1})$.

The parameters β_1 and β_2 are the ones that control the clustering (or spatial interaction) in the lattice. β_1 controls the clustering in the E-W direction and β_2 controls the N-S clustering.

The 1st-order binary scheme described by (3.2.5) is said to be isotropic if $\beta_1 = \beta_2 = \beta$ (say). Thus, a homogeneous isotropic 1st-order auto-logistic model is described by

$$P[x_{i,j} | N\text{-}N] = \frac{\exp\{[\alpha + \beta\, t_{i,j}]\, x_{i,j}\}}{1 + \exp[\alpha + \beta\, t_{i,j}]} \tag{3.2.6}$$

where $t_{i,j}$ = sum of the N-N values

$$= x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}$$

Remark 3.2.4: The number of parameters required in a binary scheme depends upon the order of the near-neighbours considered.

For Example: A second-order model involves cliques of size three and four so that the expression for the conditional probability structure is given by

$$P[x|t,t', u,u', v,v', w,w'] = \frac{\exp[x\ T]}{1 + \exp[T]}$$

where:

$$T = \alpha + \beta_1(t + t') + \beta_2(u + u') + \gamma_1(v + v') + \gamma_2(w + w')$$

$$+ \xi_1(tu + u'w + w't') + \xi_2(tv + v'u' + ut')$$

$$+ \xi_3(tw + w'u + u't') + \xi_4(tu' + uv + v't')$$

$$+ \eta(tuv + t'u'v' + tu'w' + t'uw')$$

The above scheme will be auto-logistic only if all the ξ and η parameters are zero.

3.2.3 Test of randomness in case of one-dimensional auto-logistic model:

In the class of Markov random fields given by (3.2.4) the sub-class of Markov random fields with no spatial interaction is characterized by $\beta = 0$. Consequently testing for randomness amounts to testing $\beta = 0$ against $\beta \neq 0$, indicating a spatial interaction.

In terms of notations we are interested in testing

$$(3.2.7)\quad \begin{cases} H_0: & \beta = 0 \quad \text{(i.e. no spatial interaction)} \\ H_a: & \beta \neq 0 \quad \text{(i.e. there is a spatial interaction)} \end{cases}$$

Under H_0 (3.2.4) gives

$$P_0[X_i | \text{N-N}] = \frac{e^{\alpha X_i}}{1 + e^{\alpha}}$$

which is independent of the N-N values. Thus, under H_0: $\{X_i : i \geq 1\}$ is a sequence of i.i.d. binary random variables. Consequently the joint distribution of $(X_1, \ldots, X_n)$ under H_0 is

$$f_0(X_1, \ldots, X_n) = \frac{e^{\alpha \sum_1^n X_i}}{(1 + e^{\alpha})^n}$$

implies

$$\ln f_0(X_1, \ldots, X_n) = \alpha \sum_{i=1}^{n} X_i - n \ln (1 + e^{\alpha})$$

Further under H_a (3.2.4) gives

$$P_a[X_i | \text{N-N}] = \frac{\exp[\alpha X_i + \beta X_i(X_{i-1} + X_{i+1})]}{1 + \exp[\alpha + \beta(X_{i-1} + X_{i+1})]}$$

Thus, the joint distribution of $(X_1, \ldots, X_n)$ under H_a is

$$f_a(X_1, \ldots, X_n) = \prod_{i=1}^{n} \frac{e^{\alpha X_i + \beta X_i(X_{i-1} + X_{i+1})}}{1 + e^{\alpha + \beta(X_{i-1} + X_{i+1})}}$$

implies

$$\ln f_a(X_1, \ldots, X_n) = \alpha \sum X_i + \beta \sum X_i(X_{i-1} + X_{i+1})$$
$$- \sum_1^n \ln [1 + e^{\alpha + \beta(X_{i-1} + X_{i+1})}]$$

Let $L_n(\beta)$ denote the log-likelihood function so that we have:

(3.2.8) $$L_n(0) = \alpha \sum_{i=1}^{n} X_i - n\,\ln(1 + e^{\alpha})$$

$$L_n(\beta) = \sum_{i=1}^{n} \{\alpha X_i + \beta X_i(X_{i-1} + X_{i+1})$$

$$- \ln[1 + e^{\alpha + \beta(X_{i-1} + X_{i+1})}]\}$$

i.e.,

(3.2.9) $$L_n(\beta) = \sum_{i=1}^{n} \ell_i(\beta) \quad \text{(say)}$$

where

(3.2.9)' $$\ell_i(\beta) = \alpha X_i + \beta X_i(X_{i-1} + X_{i+1}) - \ln\{1 + e^{\alpha + \beta(X_{i-1} + X_{i+1})}\}.$$

The likelihood equation is:

$$0 = \frac{\partial}{\partial\beta} L_n(\beta) = \sum_1^n \frac{\partial}{\partial\beta} \ell_i(\beta) = \sum_1^n \ell_i'(\beta)$$

$$= \sum_1^n \{\ell_i'(0) + \beta\ell_i''(0) + \frac{\beta^2}{2} \ell_i'''(\beta^*)\}$$

where $|\beta^*| < |\beta|$. (By Taylor's series expansion of $\ell_i'(\beta)$ about $\beta = 0$)

or

$$0 = \frac{1}{n} \frac{\partial}{\partial\beta} L_n(\beta) = \frac{1}{n} \sum_1^n \ell_i'(0) + \frac{\beta}{n} \sum_{i=1}^{n} \ell_i''(0)$$

$$+ \frac{\beta^2}{2n} \sum_{i=1}^{n} \ell_i'''(\beta^*)$$

(where $|\beta^*| < |\beta|$)

i.e.

(3.2.10) $$0 = B_0 + B_1\,\beta + B_2\,\frac{\beta^2}{2}\,.$$

where

(3.2.11) $$\begin{cases} B_0 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \ell_i'(0) \\[2ex] \quad = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} [X_i(X_{i-1} + X_{i+1}) - \dfrac{e^{\alpha}(X_{i-1} + X_{i+1})}{1 + e^{\alpha}}]; \\[2ex] B_1 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \ell_i''(0) \\[2ex] \quad = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} [\dfrac{-(X_{i-1} + X_{i+1})^2 e^{\alpha}}{(1 + e^{\alpha})^2}]; \\[2ex] B_2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \ell_i'''(\beta^*) \qquad (|\beta^*| < |\beta|) \\[2ex] \quad = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \{\dfrac{(X_{i-1} + X_{i+1})^3 e^{\alpha+\beta^*(X_{i-1} + X_{i+1})}[e^{\alpha+\beta^*(X_{i-1} + X_{i+1})} - 1]}{[1 + e^{\alpha+\beta^*(X_{i-1} + X_{i+1})}]^3} \end{cases}$$

Let $\hat{\beta}_n$ be the maximum likelihood estimate of β. Then using the classical theory for i.i.d. random variables

$$\hat{\beta}_n \xrightarrow{a.s.} 0 \quad \text{under } H_0.$$

So that $\forall\, \epsilon > 0$, $\forall\, \omega \in \Omega$ and sufficiently large n there exists a constant $M < \infty$ such that

$$P_0[\sqrt{n}|\hat{\beta}_n| > M] \leq \epsilon$$

This implies that it is enough to consider the behaviour of the terms B_0, B_1 and B_2 on the set $[|\hat{\beta}_n| \leq M/\sqrt{n}]$. (*).

Using $\hat{\beta}_n$ (3.2.10) gives

$$0 = B_0 + B_1 \hat{\beta}_n + B_2 \frac{\hat{\beta}_n^2}{2} , \tag{3.2.12}$$

(where β^* in the definition of B_2 is $\ni |\beta^*| < |\hat{\beta}_n|$ - being justified for each $\omega \in \Omega$)

Claim 1 under H_0 B_0 converges to zero in probability, as $n \to \infty$.

Proof: Under H_0: $\{X_i: i \geq 1\}$ is a sequence of i.i.d. binary random variables. Thus by theorem B-1 of Appendix B the sequence $\{X_1: i \geq 1\}$ is stationary and ergodic, and these random variables clearly have finite second moments. From theorem B-3 of Appendix B, it therefore follows that under H_0:

$$B_0 \xrightarrow{\text{Prob}} 2\, E(X_1 X_2) - \frac{e^\alpha}{1+e^\alpha} (EX_1 + EX_1)$$

$$= 2\, EX_1\, EX_2 - \frac{e^\alpha}{1+e^\alpha}\, 2\, EX_1$$

(by independence of the X_i's)

$$= 2(EX_1)^2 - \frac{2\, e^\alpha}{1+e^\alpha}\, EX_1 \quad (\because X_i\text{'s are id. dist.})$$

$$= \text{Zero, since } EX_1 = \frac{e^\alpha}{1+e^\alpha} \text{ under } H_0.$$

Claim 2: Under H_0 $B_1 \xrightarrow{a.s.} -k^2$ where $k^2 \geq 0$

Proof: From (3.2.11)

$$B_1 = -\frac{1}{n}\sum_{i=1}^{n}\frac{(X_{i-1}+X_{i+1})^2 e^{\alpha}}{(1+e^{\alpha})^2}$$

again by using theorem B-3 of Appendix B we have under H_0

$$B_1 \xrightarrow{a.s.} -\frac{e^{\alpha}}{(1+e^{\alpha})^2}\,[EX_1^2 + EX_1^2 + 2E(X_1X_2)]$$

$$= -\frac{e^{\alpha}}{(1+e^{\alpha})^2}\,[EX_1 + EX_1 + 2(EX_1)^2]$$

(i.i.d. property of the X_i's and the fact $X_i^2 = X_i$ for X_i binary)

$$= \frac{-e^{\alpha}}{(1+e^{\alpha})^2}\; 2\, EX_1(1 + EX_1)$$

$$= -k^2 \text{ where } k^2 = \frac{2EX_1(1+EX_1)e^{\alpha}}{(1+e^{\alpha})^2} \geq 0$$

Claim 3: Under H_0 as $n \to \infty$, B_2 is asymptotically bounded.

Proof: From (3.2.11) we have

$$B_2 = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(X_{i-1}+X_{i+1})^3 e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}[e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}-1]}{[1+e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}]^3}\right\}$$

(where β^* is such that $|\beta^*| < |\hat{\beta}_n|$)

$$\leq \max_{1\leq i\leq n}\frac{(X_{i-1}+X_{i+1})^3 e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}[e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}-1]}{[1+e^{\alpha+\beta^*(X_{i-1}+X_{i+1})}]^3}$$

$$\leq \frac{2^3\, e^{\alpha+2\beta^*}\,[e^{\alpha+2\beta^*}-1]}{(1+e^{\alpha})^3}$$

(since X_i's are binary $\min(X_{i-1}+X_{i+1}) = 0$ and $\max(X_{i-1}+X_{i+1}) = 2$)

$$\leq \frac{2^3 \, e^{\alpha + \frac{2M}{\sqrt{n}}} [e^{\alpha + \frac{2M}{\sqrt{n}}} - 1]}{(1 + e^{\alpha})^3} \qquad \text{(by *)}$$

which is bounded as $n \to \infty$.

Consequently it follows that as $n \to \infty$ under H_0 B_2 is asymptotically bounded. (3.2.12) can be rewritten as

$$\sqrt{n}\,\hat{\beta}_n = \frac{\sqrt{n}\,B_0}{-B_1 - B_2 \frac{\hat{\beta}_n}{2}}$$

i.e.

$$\sqrt{n}\,\hat{\beta}_n = \frac{\sqrt{n}\,B_0/k^2}{-\frac{B_1}{k^2} - B_2 \frac{\hat{\beta}_n}{2k^2}} \tag{3.2.13}$$

using claims 2 and 3 and the fact that $\hat{\beta}_n \overset{P}{\to} 0$ under H_0, it follows that the denominator of (3.2.13) converges to 1 in probability as $n \to \infty$.

Thus the asymptotic distribution of $\sqrt{n}\,\hat{\beta}_n$ depends upon the asymptotic distribution of $\sqrt{n}\,B_0$.

<u>Asymptotic Distribution of $\sqrt{n}\,B_0$ under H_0:</u>

Using (3.2.11) we have

$$\sqrt{n}\,B_0 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \{X_i(X_{i-1} + X_{i+1}) - \frac{e^{\alpha}(X_{i-1} + X_{i+1})}{1 + e^{\alpha}}\}$$

i.e.

$$\sqrt{n}\,B_0 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i \quad \text{(say)}$$

where $\xi_i = X_i(X_{i-1} + X_{i+1}) - \dfrac{e^{\alpha}(X_{i-1} + X_{i+1})}{1 + e^{\alpha}}$; $i \geq 1$.

Under H_0: $\{X_i: i \geq 1\}$ is a sequence of i.i.d. binary random variables. Therefore we have

$$E\,\xi_i = 0 \quad \forall\, i$$

and that $\xi_i, \xi_{i+3}, \ldots$ are independent so that $\{\xi_i: i \geq 1\}$ is a 2-dependent stationary process with mean zero. Consequently, by theorem B-2 of Appendix B it follows that $\sqrt{n}\,B_0 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i$ has a limiting normal distribution with mean zero and variance

$$\sigma_\xi^2 = E\,\xi_1^2 + 2E\,\xi_1\xi_2 + 2E\,\xi_1\xi_3 = \frac{e^{2\alpha}(3 + e^{\alpha})}{(1 + e^{\alpha})^4}$$

Hence it follows from (3.2.13) that $\sqrt{n}\,\hat{\beta}_n$ has a limiting normal distribution with mean zero and variance σ_ξ^2/k^4.

<u>Asymptotic distribution of the log-likelihood ratio $\Lambda_n(\hat{\beta}_n,0) = L_n(\hat{\beta}_n) - L_n(0)$, under H_0:</u>

Expanding $L_n(0)$ by Taylor's deries expansion about $\hat{\beta}_n (\forall\, \omega \in \Omega)$ we get

$$L_n(0) = L_n(\hat{\beta}_n) + (0-\hat{\beta}_n)L_n'(\hat{\beta}_n) + \frac{(0-\hat{\beta}_n)^2}{2}\, L_n''(\beta^*)$$

where $|\beta^*| < |\hat{\beta}_n|$

Hence: $\Lambda_n(\hat{\beta}_n,0) = L_n(\hat{\beta}_n) - L_n(0)$

$$= \hat{\beta}_n L_n'(\hat{\beta}_n) - \frac{\hat{\beta}_n^2}{2}\, L_n''(\beta^*)$$

i.e.,

$$(3.2.14)\quad \Lambda_n(\hat{\beta}_n,0) = \hat{\beta}_n L_n'(\hat{\beta}_n) - \frac{\hat{\beta}_n^{\,2}}{2}[L_n''(\beta^*) - L_n''(0)] - \frac{\hat{\beta}_n^{\,2}}{2} L_n''(0)$$

Since $L_n'(\hat{\beta}_n) = 0$ and under H_0 $\hat{\beta}_n \xrightarrow{a.s.} 0$ it follows that the 1st term on the r.h.s. of (3.2.14) converges to zero in probability under H_0.

$$\text{3rd term} = -\frac{\hat{\beta}_n^{\,2}}{2} L_n''(0)$$

$$= \left(\frac{\sqrt{n}\hat{\beta}_n}{\sigma_\xi/k^2}\right)^2 \frac{\sigma_\xi^{\,2}}{2k^4}\left(-\frac{1}{n}L_n''(0)\right)$$

From the preceding pages $\dfrac{\sqrt{n}\hat{\beta}_n}{\sigma_\xi/k^2} \xrightarrow{\mathcal{D}} N(0,1)$ and by claim 2 $-\frac{1}{n}L_n''(0) \xrightarrow{a.s.} k^2$ under H_0. Consequently the 3rd term on the r.h.s. of (3.2.14) converges in distribution to $(N(0,1))^2\,\sigma_\xi^{\,2}/2k^2$.

$$\text{2nd term} = \frac{\hat{\beta}_n^{\,2}}{2}[L_n''(0) - L_n''(\beta^*)]$$

$$= (\sqrt{n}\hat{\beta}_n)^2 \frac{1}{2n}[L_n''(0) - L_n''(\beta^*)]$$

$$= (\sqrt{n}\hat{\beta}_n)^2 \frac{\epsilon_n}{2}\ .$$

By following an argument analogous to that of Borwankar et at [1971] it follows that $|\epsilon_n| \to 0$ as $n \to \infty$ (For a proof see Appendix B). Hence under H_0 the 2nd term also converges to zero in probability. Thus: under H_0, $\Lambda_n(\hat{\beta}_n,0)$, the log-likelihood ratio, follows a central chi-squared distribution, and the associated degrees of freedom is 1.

3.2.4 Test of randomness for a homogeneous 1st-order (N-N) isotropic auto-logistic model:

Following remark 3.2.3, a homogeneous 1st-order isotropic auto-logistic model is given by

$$P[X_{ij}|N-N] = \frac{e^{[\alpha+\beta\, t_{ij}]x_{i,j}}}{1+e^{\alpha+\beta t_{ij}}}$$

where $t_{i,j} = x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}$.

The hypothesis of interest are:

$$\begin{cases} H_0: & \beta = 0 \quad \text{(no spatial interaction)} \\ H_a: & \beta \neq 0 \quad \text{(there is a spatial interaction)} \end{cases}$$

The procedure to get the asymptotic distribution of the log=likelihood ratio $\Lambda_n(\hat{\beta}_n,0) = L_n(\hat{\beta}_n) - L_n(0)$, ($\hat{\beta}_n$ being the maximum likelihood estimator of β and $L_n(\beta)$, the log-likelihood function) under H_0 is analogous to the one given in section 3.2.3 for the one-dimensional case. So we only sketch the main lines of proof.

$$L_n(\beta) = \sum_{i,j} \{\alpha x_{i,j} + \beta t_{i,j} x_{i,j} - \ell n[1 + e^{\alpha+\ \beta t_{i,j}}]\}$$

$$= \sum_{i,j} \ell_{i,j}(\beta)$$

where $\ell_{i,j}(\beta) = \alpha\ x_{i,j} + \beta\ t_{i,j}\ x_{i,j} - \ell n\ [1 + e^{\alpha+\beta\ t_{i,j}}]$

Likelihood equation is:

$$0 = \frac{\partial}{\partial\beta} L_n(\beta) = \sum_{i,j} \frac{\partial}{\partial\beta} \ell_{i,j}(\beta)$$

or

$$0 = \frac{1}{n^2}\frac{\partial}{\partial\beta} L_n(\beta) = \frac{1}{n^2} \sum_{i,j} \frac{\partial}{\partial\beta} \ell_{i,j}(\beta)$$

If $\hat{\beta}_n$ denotes the maximum likelihood estimator of β we have then

$$0 = \frac{1}{n^2} \sum_{i,j} \ell'_{i,j}(\hat{\beta}_n)$$

$$= \frac{1}{n^2} \sum_{i,j} [\ell'_{i,j}(0) + \hat{\beta}_n \ell''_{i,j}(0) + \frac{\hat{\beta}_n^2}{2} \ell'''_{i,j}(\beta^*)]$$

(where $|\beta^*| < |\hat{\beta}_n|$ being justified $\forall\ \omega \in \Omega$.)

i.e.

$$0 = B_0 + B_1 \hat{\beta}_n + B_2 \frac{\hat{\beta}_n^2}{2} \tag{3.2.15}$$

where

$$\left\{ \begin{aligned} B_0 &= \frac{1}{n^2} \sum_{i,j} \ell'_{i,j}(0) \\ &= \frac{1}{n^2} \sum_{i,j} \{t_{i,j} x_{i,j} - \frac{e^{\alpha}\, t_{i,j}}{1 + e^{\alpha}}\} \\ B_1 &= -\frac{1}{n^2} \sum_{i,j} \{t_{i,j}^2\ e^{\alpha}/(1 + e^{\alpha})^2\} \\ B_2 &= \frac{1}{n^2} \sum_{i,j} \{ \frac{t_{i,j}^3\ e^{\alpha+\beta^* t_{i,j}} [e^{\alpha+\beta^* t_{i,j}} - 1]}{(1 + e^{\alpha+\beta^* t_{i,j}})^3} \} \end{aligned} \right. \tag{3.2.16}$$

where $t_{i,j} = x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}$ and $|\beta^*| < |\hat{\beta}_n|$.

As dealt with in section 3.2.3 we have under H_0:

$$B_0 \xrightarrow{a.s.} 0 \quad \text{as} \quad n \to \infty$$

$$B_1 \xrightarrow{a.s.} -k_1^2 \quad (k_1^2 \geq 0) \quad \text{as} \quad n \to \infty$$

and B_2 is asymptotically bounded as $n \to \infty$.

Rewriting (3.2.15) we get:

$$\hat{\beta}_n = \frac{B_0}{-B_1 - \frac{\hat{\beta}_n}{2} B_2}$$

implies

$$n\hat{\beta}_n = \frac{nB_0/k_1^2}{-\frac{B_1}{k_1^2} - \frac{\hat{\beta}_n}{2k_1^2} B_2}$$

As before, the denominator converges to 1 under H_0 so that the asymptotic distribution of $n\hat{\beta}_n$ depends upon that of nB_0.

Further:

$$nB_0 = \frac{1}{n} \sum_{i,j} \{t_{i,j}x_{i,j} - \frac{e^{\alpha}t_{i,j}}{1 + e^{\alpha}}\}$$

$$= \frac{1}{n} \sum_{i,j} \xi_{i,j}$$

where

$$\xi_{i,j} = t_{i,j}x_{i,j} - \frac{e^{\alpha}t_{i,j}}{1 + e^{\alpha}}$$

$$= [x_{i-1,j}x_{i,j} + x_{i+1,j}x_{i,j}$$

$$+ x_{i,j-1}x_{i,j} + x_{i,j+1}x_{i,j}]$$

$$- \frac{e^{\alpha}}{1 + e^{\alpha}} [x_{i-1,j} + x_{i+1,j} + x_{i,j-1} + x_{i,j+1}]$$

is under H_0 a 4-dependent stationary process with mean zero. An extension of theorem B-3 of appendix implies that nB_0 has a limiting normal distribution with mean zero and some variance σ_ξ^2 and consequently $n\hat{\beta}_n$ is asymptotically normal with mean zero and variance σ_ξ^2/k_1^4.

Finally following exact similar lines of proof as in section 3.2.3, it follows that $\Lambda_n(\hat{\beta}_n,0) = L_n(\hat{\beta}_n) - L_n(0)$ has under H_0 a central chi-square distribution and the associated degrees of freedom is 3.

Remark 3.2.5: [Besag - 1974] considers the above problem and states the chi-square behaviour without giving proper justification. Also it was seen that the coding method of estimation as suggested by Besag is not necessary in establishing the chi-square behaviour. For further details on coding method of estimation one is referred to Besag's paper.

CHAPTER IV

POWER UNDER A SPECIFIC ALTERNATIVE

In Chapter III, we looked at some of the estimators of λ, the intensity of the spatial process, and the asymptotic properties of these estimators. It was also seen that in a particular subclass of auto-binary schemes the test statistic for testing randomness (no spatial interaction) has a chi-square behaviour under the null hypothesis of complete randomness. However, the value of a test statistic is increased if one can discuss the power of the test to detect departure from randomness.

In this chapter, we try to look at the distribution of the test statistic under a specific alternative. We shall confine our attention to the subclass of auto-binary schemes where the parameter describing the spatial interaction has a specified form under the alternative. The problem has been approached using ideas on contiguity (discussed in Section 4.1). It has been shown that in our particular formulation the classical techniques of contiguity fail. Even though the measures are shown to be contiguous using the basic principles of contiguity, the distribution of the test statistic under the alternative does not seem to have an easy tractable form. A conceivable conjecture is that it depends upon a non-central chi-square distribution (Section 4.2).

4.1 Contiguity and its characterizations

Our exposition follows Roussas [1972].

The concept of contiguity was first introduced by Professor Le Cam as a measure of 'nearness' of sequences of probability measures. It plays an important role in the study of asymptotic theory in deriving asymptotic properties of the tests under much less assumptions.

Definition 4.1.1: Let $\{(X,A_n)\}$ be a sequence of measurable spaces and $\{P_n\}$ and $\{Q_n\}$ be two sequences of probability measures defined on (X,A_n). The sequence $\{Q_n\}$ is said to be contiguous with respect to $\{P_n\}$ if and only if $\forall A_n \in A_n$

(4.1.1.) $$[P_n(A_n) \to 0 \quad \text{implies} \quad Q_n(A_n) \to 0]$$

In such a case, we also say that the densities q_n are contiguous to the densities p_n where p_n and q_n are respectively the densities of P_n and Q_n with respect to some dominating σ-finite measure.

Remark 4.1.1: Contiguity implies that any sequence of random variables converging to zero in P_n-probability converges to zero in Q_n-probability.

Definition 4.1.2: The sequence P_n of probability measures is said to be relatively compact if for every subsequence $\{n'\}$ of $\{n\}$ there exists a further subsequence $\{n''\}$ of $\{n'\}$ such that $P_{n''}$ converges weakly to a probability measure P (definition 1.4.1).

Alternative characterizations of contiguity:

Let p_n and q_n be the densities respectively of P_n and Q_n with respect to some dominating σ-finite measure.

Define the log-likelihood ratio Λ_n as follows:

(4.1.2) $$\Lambda_n = \begin{cases} \log(q_n | p_n) & \text{on } \{p_n q_n > 0\} \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

For each determination of Λ_n defined by (4.1.2) let

(4.1.3) $$\begin{aligned} L_n &= L[\Lambda_n | P_n] \\ L_n' &= L[\Lambda_n | Q_n] \end{aligned}$$

Therorem 4.1.1: (Roussas [1972] pp 11-14)

The following three statements are equivalent:

(i) $\{P_n\}$ and $\{Q_n\}$ are contiguous;

(ii) $\{L_n\}$ and $\{L_n'\}$ are relatively compact (for each determination of Λ_n);

(iii) $\{L_n\}$ is relatively compact and if F is the limiting distribution function and $X \sim F$ then

$$\int e^x \, dF(x) = 1.$$

Remarks 4.1.2: (i) L_1-norm convergence implies contiguity

i.e., $$\|P_n - Q_n\|_1 \to 0$$

implies $\{P_n\}$ and $\{Q_n\}$ are contiguous where $\|P_n - Q_n\|_1$ is defined by

$$\|P_n - Q_n\|_1 = 2 \sup_A |P_n(A) - Q_n(A)|.$$

(For a proof see Roussas [1972], p. 9)

(ii) converse of (i) is not true

i.e., contiguity does not imply L_1-norm convergence.

<u>Example</u>: Let $(X,A_n) = (R,B)$. Take $P_n \equiv N(\mu_n,1)$ and $Q_n = N(\mu_n{}',1)$ where $\mu_n \to \mu$ and $\mu_n' \to \mu'$ and $\mu,\mu'(\mu \neq \mu')$ are both finite. It can easily be seen that $\Lambda_n = (\mu_n' - \mu_n) X + \frac{\mu_n^2 - \mu_n'^2}{2}$

so that

$$L_n = L(\Lambda_n|P_n) = N[\frac{-(\mu_n' - \mu_n)^2}{2} \quad , \quad (\mu_n' - \mu_n)^2]$$

$$L_n' = L(\Lambda_n|Q_n) = N[\frac{(\mu_n' - \mu_n)^2}{2} \quad , \quad (\mu_n' - \mu_n)^2].$$

clearly $\{L_n\}$ and $\{L_n'\}$ are relatively compact and consequently by theorem 4.1.1 $\{P_n\}$ and $\{Q_n\}$ are contiguous. However $\|P_n - Q_n\|_1 \not\to 0$.

(iii) Contiguity does not imply mutual absolute continuity:

<u>Example</u>: Let $(X,A_n) = (R,B)$

Take $P_n = U(-\frac{1}{n},1)$, the uniform measure on $(-\frac{1}{n}, 1)$

and $Q_n = U(0,1 + \frac{1}{n})$ the uniform measure on $(0, 1 + \frac{1}{n})$

Further $p_n(x) = \frac{dP_n(x)}{d\nu} = \frac{n}{n+1} \quad , \quad -\frac{1}{n} < x < 1$

and $\quad q_n(x) = \frac{dQ_n(x)}{d\nu} = \frac{n}{n+1} \quad 0 < x < 1 + \frac{1}{n}$.

ν being the Lebesgue measure.

Then $\|P_n - Q_n\|_1 = \frac{2}{n+1} \to 0$ as $n \to \infty$ so that by (ii) above $\{P_n\}$ and $\{Q_n\}$ are contiguous.

However $\{P_n\}$ and $\{Q_n\}$ are clearly not mutually absolutely continuous.

(iv) Mutually absolutely continuous need not imply contiguity:

Example: Let $(\mathcal{X},\mathcal{A}_n) = (\mathcal{R},\mathcal{B})$

Take $P_n \equiv N(\mu_n,1)$ $Q_n \equiv N(\mu_n',1)$ where $\mu_n \to -\infty$, and $\mu_n' \to +\infty$ Then $\{P_n\}$ and $\{Q_n\}$ are mutually absolutely continuous but not contiguous for taking $A_n = (\mu_n - 1, \mu_n + 1)$ we see that $P_n(A_n) = .68 \not\to 0$ but $Q_n(A_n) \to 0$.

(v) The above (i) - (iv) imply that contiguity is weaker than the L_1-norm convergence and is distinct from the mutual absolute continuity notion.

4.1.1 Some results following from contiguity:

In this section we shall discuss some important consequences of the notion of contiguity and their use in statistical applications.

Lemma 4.1.1: [Roussas - 1972, p. 15]

Any one of the three equivalent statements of theorem 4.1.1 implies that $P_n(B_n) \to 1$ and $Q_n(B_n) \to 1$ where $B_n = \{p_n q_n > 0\}$.

Remark 4.1.3: Lemma 4.1.1 says that contiguous measures P_n and Q_n eventually rest on B_n i.e. eventually are mutually absolutely continuous. Consequently if $\{P_n\}$ and $\{Q_n\}$ are contiguous, one may assume without loss of generality that P_n and Q_n are mutually absolutely continuous for all sufficiently large n. Under this assumption the log-likelihood ratio

$$\Lambda_n = \log(dQ_n/dP_n)$$

is well-defined a.s. (P_n), (Q_n).

Define

$$
(4.1.4) \qquad \begin{cases} \mathcal{L}_n^* = \mathcal{L}[(\Lambda_n, T_n)|P_n] \\ \\ \mathcal{L}_n'^* = \mathcal{L}[(\Lambda_n, T_n)|Q_n] \end{cases}
$$

where $\{T_n\}$ is a sequence of k-dimensional random vectors such that T_n is $\mathcal{A}_n$-measurable.

Theorem 4.1.2: [Roussas - 1972, p. 34]

Suppose $\{P_n\}$ and $\{Q_n\}$ are contiguous and $\mathcal{L}_n^*$ and $\mathcal{L}_n'^*$ are defined by (4.1.4). Further assume that $\mathcal{L}_n^* \Rightarrow \mathcal{L}^*$, a probability measure.

Then $\qquad \mathcal{L}_n'^* \Rightarrow \mathcal{L}'^*$ where

$$
\frac{d\mathcal{L}'^*}{d\mathcal{L}^*} = \exp(x)
$$

Corollary 4.1.1: [Roussas - 1972, p. 35]

If $\mathcal{L}(\Lambda_n|P_n) \Rightarrow N(\mu,\sigma^2)$ then $\mu = -\frac{1}{2}\sigma^2$.

Corollary 4.1.2: If $\mathcal{L}(\Lambda_n|P_n) \Rightarrow N(-\frac{1}{2}\sigma^2,\sigma^2)$ then $\mathcal{L}(\Lambda_n|Q_n) \Rightarrow N(\frac{1}{2}\sigma^2,\sigma^2)$.

4.1.2 Interpretation of contiguity in simple versus simple hypotheses testing:

Consider a sequence $\{p_n,q_n\}$ of simple hypothesis p_n against simple alternatives q_n defined on measurable spaces $(\mathcal{X}_n, \mathcal{A}_n)$ respectively.

According to Neyman-Pearson lemma, for any event A_n in $\mathcal{A}_n$, there exists a function Φ_n and an integer k_n: $0 < k_n < \infty$ such that

$$(4.1.5) \qquad \Phi_n = \begin{cases} 1 & \text{if} \quad q_n > k_n p_n \\ \xi(0 \leq \xi \leq 1) & \text{if} \quad q_n = k_n p_n \\ 0 & \text{if} \quad q_n < k_n p_n \end{cases}$$

and that

$$P_n(A_n) = \int \Phi_n \, dP_n$$

and

$$Q_n(A_n) \leq \int \Phi_n \, dQ_n$$

Thus contiguity (definition 4.1.1) will follow if we can show that

$$[\int \Phi_n \, dP_n \to 0] \quad \text{implies} \quad [\int \Phi_n \, dQ_n \to 0]$$

for critical functions of the type (4.1.5).

Remark 4.1.4: Using the equivalence of the statements (i) and (iii) of theorem 4.1.1, it can easily be observed that if Λ_n, the log-likelihood ratio, is asymptotically normal $(-\frac{1}{2}\sigma^2, \sigma^2)$ under P_n, then the densities q_n and p_n are contiguous (For a proof see Hajek & Sidek 1967, pp 203-205).

Suppose we have an experiment $\{(X_n, A_n, \Theta): \Theta \subseteq R^k\}$ and are interested in testing

$$(4.1.6) \quad \begin{cases} H_0: \ \theta = \theta_0 \\ H_{a,n}: \ \theta = \theta_n = \theta_0 + \delta_n h_n \end{cases}$$

where $\delta_n \downarrow 0$ and $\{h_n\}$ is a bounded sequence in R^k such that $h_n \to h \in R^k$.

The following proposition says that under certain conditions the probability measures P_{n,θ_0} and P_{n,θ_n} are contiguous.

<u>Proposition 4.1.1</u>: Consider θ_0 and $\theta_n = \theta_0 + \delta_n h_n$ with $\delta_n \to 0$ and $\{h_n\}$ bounded and $h_n \to h$. $(h_n, h \in R^k)$.
Suppose there exists an A_n-measurable function $T_n(\theta)$ and a positive definite covariance matrix Γ_θ such that

$$(4.1.7) \quad \begin{cases} a) \ \Lambda_n(\theta_n, \theta_0) - h' \, T_n(\theta_0) + \tfrac{1}{2} h' \, \Gamma_{\theta_0} h \to 0 \ (P_{n,\theta_0}) \\ b) \ L[T_n(\theta_0) | P_{n,\theta_0}] \to N(\overline{0}, \ \Gamma_{\theta_0}) \end{cases}$$

then P_{n,θ_0} and P_{n,θ_n} are contiguous where

$$\Lambda_n(\theta_n, \theta_0) = \log \frac{dP_{n,\theta_n}}{dP_{n,\theta_0}}$$

<u>Proof</u>: Assumption (4.1.7) implies that

$$L[\Lambda_n(\theta_n, \theta_0) | P_{n,\theta_0}] \to N(-\tfrac{1}{2} h' \, \Gamma_{\theta_0} h, \ h' \, \Gamma_{\theta_0} h)$$

By remark (4.1.1) it then follows that P_{n,θ_0} and P_{n,θ_n} are contiguous.

Remark 4.1.5: (i) Under the conditions of proposition 4.1.1, it follows that

$$L(\Lambda_n(\theta_n,\theta_0)|P_{n,\theta_n}) \to N(\tfrac{1}{2}\, h'\, \Gamma_{\theta_0} h,\ h'\, \Gamma_{\theta_0} h)$$

(ii) Condition (4.1.7) is known as the locally asymptotically normal (LAN) conditions.

Example 4.1.1:

Take $P_{n,\theta_0} \equiv N(\theta_0,1)$

and $P_{n,\theta_n} = N(\theta_n,1)$ with $\theta_n = \theta_0 + n^{-\frac{1}{2}}h$.

One can easily see that

$$\Lambda_n(\theta_n,\theta_0) = (\theta_n-\theta_0)\ X + \frac{\theta_0^2 - \theta_n^2}{2}$$

Take $T_n(\theta) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \theta)$ and $\Gamma_\theta \equiv 1$.

It can be verified that ℓ.h.s of (4.1.7) (a) is identically zero and the ℓ.h.s. of (4.1.7) (b) is exactly normal.

Thus, by proposition 4.1.1 the two probability measures P_{n,θ_0} and P_{n,θ_n} where $\theta_n = \theta_0 + n^{-\frac{1}{2}}h$ (h bounded) are contiguous.

Remark 4.1.5: We describe below a typical problem that comes up in statistical applications.

Suppose $\{P_n\}$ and $\{Q_n\}$ are two sequences of probability measures and that Q_n depends on an h in a specified way. Further suppose that assumption (4.1.7) holds. The problem is that of finding the asymptotic distribution of T_n under Q_n.

What one does then is to first derive the asymptotic destribution of (Λ_n, T_n) under P_n and use the contiguity of P_n and Q_n to get the asymptotic distribution of (Λ_n, T_n) under Q_n. From this, the desired asymptotic distribution of T_n under Q_n will follow.

4.2 Power of the test:

We are back in the particular subclass of auto-binary spatial schemes. In Section 3.2, we discussed the asymptotic behaviour of the log-likelihood ratio under the null hypothesis of complete randomness. In this section, we attempt to discuss the asymptotic distribution of the log-likelihood ratio under a specific alternative, $H_{a,n}$: $\beta = \beta_n = n^{-\frac{1}{2}}h$ (h being bounded). Thus we have a set of hypothesis

$$\begin{cases} H_0: & \beta = 0 \\ H_{a,n}: & \beta = \beta_n = n^{-\frac{1}{2}}h \text{ (h bounded)} \end{cases}$$

So that $\beta_n \to 0$ as $n \to \infty$. Without any loss of generality one may take $h \equiv 1$ so that the hypotheses are

$$\text{(4.2.1)} \qquad \begin{cases} H_0: & \beta = 0 \\ H_{a,n}: & \beta = \beta_n = n^{-\frac{1}{2}} \to 0 \quad \text{as} \quad n \to \infty \end{cases}$$

where β is the parameter describing the spatial interaction in the auto-binary spatial models (both one-dimensional model of Section 3.2.2 and two-dimensional isotropic model of Section 3.2.4).

Following remark 4.1.5, in order to determine the asymptotic distribution under the alternative, knowing the asymptotic distribution under the null the first step is to show the contiguity of measures under H_0 and $H_{a,n}$. Also, proposition 4.1.1 says that contiguity will follow if the LAN conditions given by (4.1.7) can be verified for our model. However, unfortunately the LAN conditions are not satisfied in our formulation (see Appendix C) so that the classical techniques of contiguity fail.

4.2.1 Contiguity of measures under the hypotheses defined by (4.2.1)

Since the classical techniques of contiguity fail in our formulation we approach the problem through basic principles of contiguity (definition 4.1.1). Let $(X,\mathcal{A}_n)$ be a measurable space and $\{P_\beta\}$ be a sequence of probability measures defined on $(X,\mathcal{A}_n)$. Let $A_n \in \mathcal{A}_n$. Then we would like to show that

$$[P_0(A_n) \to 0] \text{ implies } [P_{\beta_n}(A_n) \to 0]$$

<u>One-dimensional model</u>:

Let $A_n \in \mathcal{A}_n$ and $P_0(A_n) \to 0$ would like to show that

$$P_{\beta_n}(A_n) \to 0 \quad \text{as} \quad n \to \infty$$

The conditional probability distribution of an auto-binary spatial scheme is given by

$$\frac{dP_\beta}{d\mu}[X_i|N\text{-}N] = \frac{e^{[\alpha + \beta(X_{i-1} + X_{i+1})]X_i}}{1 + e^{\alpha + \beta(X_{i-1} + X_{i+1})}}$$

(μ being some dominating σ-finite measure)

Consequently it follows that

$$\frac{dP_{\beta_n}[X_i|N-N]}{dP_0[X_i|N-N]} = \frac{e^{\beta_n(X_{i-1}+X_{i+1})X_i}(1+e^{\alpha})}{[1+e^{\alpha+\beta_n(X_{i-1}+X_{i+1})}]} \tag{4.2.2}$$

where X_i's take values 0 or 1 and $\beta_n = n^{-\frac{1}{2}}$ converges to zero as $n \to \infty$.

Since X_i's take only 0-1 values, $X_i(X_{i-1} + X_{i+1})$ is at most 2.

Further, the likelihood ratio is given by

$$\ln(\beta_n, 0) = \prod_{i=1}^{n} \frac{dP_{\beta_n}[X_i|N-N]}{dP_0[X_i|N-N]}$$

i.e.,

$$\ln(\beta_n, 0) = \prod_{i=1}^{n} \{\frac{e^{\beta_n X_i(X_{i-1}+X_{i+1})}}{1+e^{\alpha+\beta_n(X_{i-1}+X_{i+1})}}) \quad (1+e^{\alpha})\} \tag{4.2.3}$$

Now,

$$\begin{aligned}
P_{\beta_n}(A_n) &= \int_{A_n} \frac{dP_{\beta_n}}{dP_0} \quad dP_0 \\
&= \int \chi_{A_n} \frac{dP_{\beta_n}}{dP_0} \quad dP_0 \\
&= \int \chi_{A_n} \prod_{i=1}^{n} \frac{e^{\beta_n X_i(X_{i-1}+X_{i+1})}}{1+e^{\alpha+\beta_n(X_{i-1}+X_{i+1})}} \quad (1+e^{\alpha})dP_0 \\
&\leq (1+e^{\alpha}) \int \chi_{A_n} \prod_{i=1}^{n} \frac{e^{2\beta_n}}{1+e^{\alpha}} dP_0 \\
&= (1+e^{\alpha}) \int \chi_{A_n} (\frac{e^{2\beta_n}}{1+e^{\alpha}})^n dP_0
\end{aligned}$$

i.e.

$$(4.2.4) \qquad P_{\beta_n}(A_n) \leq (1 + e^{\alpha}) \left(\frac{e^{2\beta_n}}{1+e^{\alpha}} \right)^n P_0(A_n)$$

<u>Claim</u> $\left(\frac{e^{2\beta_n}}{1+e^{\alpha}} \right)^n$ (for $\beta_n = n^{-\frac{1}{2}}$) $\to 0$ as $n \to \infty$.

for any $\alpha > 0$:

$$\text{Now} \quad \left(\frac{e^{2\beta_n}}{1+e^{\alpha}} \right)^n \leq \frac{e^{2n^{-\frac{1}{2}}n}}{e^{\alpha n}}$$

$$= \frac{e^{2n^{\frac{1}{2}}}}{e^{\alpha n}}$$

$$= \frac{1}{e^{\alpha n - 2n^{\frac{1}{2}}}}$$

$$\to 0 \text{ as } n \to \infty \quad \text{for } \alpha > 0.$$

Since $P_0(A_n) \to 0$ (by hypothesis), therefore it follows from (4.2.4) that

$$P_{\beta_n}(A_n) \to 0$$

thus establishing the contiguity of measures under H_0 and $H_{a,n}$ defined by (4.2.1).

<u>Remark 4.2.1</u>: In case of a 2-dimensional isotropic auto-binary model also, the lines of proof (as given above) give the contiguity of measures under H_0 and $H_{a,n}$, defined by (4.2.1).

<u>Remark 4.2.2</u>: We have not yet managed to establish the asymptotic distribution of the log-likelihood under the alternative. A good conjecture is that it is a non-central chi-square distribution, which we hope to establish in the near future.

APPENDIX A

As in Section 1.1, let S denote the d-dimensional Euclidean space and $\mathcal{B}(S)$ the family of borel subsets of S. Let Ω_S be the family of point processes $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$ that satisfy conditions a), b) and c) of definition 1.1.1.

Let $\mathcal{D}$ be a countable subfamily of $\mathcal{B}(S)$ that generates $\mathcal{B}(S)$ and $F_{\mathcal{D}}$ denote the field generated by elements of $\mathcal{D}$.

<u>Definition A-1</u>: Finite dimensional distributions generated by a point process $\xi(A,\omega)$ are the distributions

$$\text{Prob}\{\xi(A_1,\omega) = r_1, \ldots, \xi(A_m,\omega) = r_m\}$$

where $A_1,\ldots,A_m \in \mathcal{B}(S)$ and $r_1,\ldots,r_m$ are non-negative integers. Let

$$q(A_1,\ldots,A_m; r_1,\ldots,r_m) = \text{Prob}\{\xi(A_1,\omega) = r_1,\ldots,\xi(A_m,\omega) = r_m\}$$

denote the finite dimensional distributions generated by a point process $\xi(A,\omega)$.

<u>Definition A-2</u>: Let $A_1,\ldots,A_m$ be sets in $F_{\mathcal{D}}$. Then a set in Ω_S determined by conditions on $\xi(A_1,\omega),\ldots,\xi(A_m,\omega)$ is called a cylinder set in Ω_S.

i.e. a set of point processes determined by its finite dimensional distributions is a cylinder set in Ω_S.

Let C be the family of cylinder sets and C^* be the borel extension of C.

Let Ω' be the family of non-negative integer valued set functions $\xi(A,\omega)$ for A in $F_{\mathcal{D}}$.

Let $C(\Omega')$ be the family of cylinder sets in Ω' and $C^*(\Omega')$ its borel extension.

Similarly define Ω'' to be the family of those set functions in Ω' that satisfy b(i) for $A_i \in F_{\mathcal{D}}$ namely $\xi(\bigcup_{i=1}^{m} A_i,\omega) = \sum_1^m \xi(A_i,\omega)$ a.e. for $A_1,\ldots,A_m$ disjoint sets in $F_{\mathcal{D}}$ and Ω''' to be the family of those set functions of Ω' that satisfy both b(i) and b(ii) for $A_i \in F_{\mathcal{D}}$.

namely

$$\xi(\bigcup_1^m A_i,\omega) = \sum_1^m \xi(A_i,\omega),\ A_i \in F_{\mathcal{D}} \text{ and disjoint}$$

and $$\xi(A_n,\omega) \to 0 \text{ for } A_1 \supset A_2 \supset \ldots \text{ in } F_{\mathcal{D}} \text{ such that } \bigcap_n A_n = \phi$$

Similar to $C(\Omega')$ and $C^*(\Omega')$ we define $C(\Omega'')$, $C^*(\Omega'')$, $C(\Omega''')$ and $C^*(\Omega''')$.

<u>Converse implication of theorem 1.1.1:</u>

Given a point process $\{\xi(A,\omega): A \in \mathcal{B}(S)\}$ that satisfies conditions a), b) and c) of definition 1.1.1 we need to show that ξ is an $(M_+(S), \mathcal{M}(S))$ - valued random variable.

i.e., Given the finite dimensional distributions $q(A_1,\ldots,A_m; r_1,\ldots,r_m)$

there exists a unique probability measure Q that is determined by these q-functions.

Conditions A-1: (n is any positive integer; $A_1,\ldots,A_m$ are sets in $\mathcal{B}(S)$ and $r_1,\ldots,r_m$ are non-negative integers):

(i) $q(A_1,\ldots,A_m;\ r_1,\ldots,r_m)$ is a probability distriubtion on m-tuples of non-negative integers $r_1,\ldots,r_m$.

Also, $q(A_1,A_2;\ r_1,r_2) = q(A_2,A_1;\ r_2,r_1)$.

(ii) The functions q are "consistent" i.e. for example

$$\sum_{r_2=0} q(A_1,A_2;\ r_1,r_2) = q(A_1,r_1).$$

(iii) If $A_1,A_2,\ldots,A_m$ are disjoint sets, and $A = A_1 \cup A_2 \cup \ldots \cup A_m$ then $q(A,A_1,\ldots,A_m;\ r_1,r_1,\ldots,r_m) = 0$ unless $r = r_1 + \ldots + r_m$ and

$$q(A,A_1,\ldots,A_m;\ r_1 +\ldots+ r_m,r_1,\ldots,r_m) = q(A_1,\ldots,A_m;\ r_1,\ldots,r_m)$$

i.e. $$\mathrm{Prob}\{\xi(A,\omega) = \sum_1^m \xi(A_i,\omega)\} = 1.$$

(Corresponds to b(i) of definition 1.1.1)

(iv) If $A_1 \supset A_2 \supset \ldots$ such that $\bigcap_n A_n = \phi$, then

$\lim_n q(A_n;\ 0) = 1$

i.e. $\mathrm{Prob}\ \{\xi(A_n,\omega) \to 0\} = 1$

(Suggested by condition b(ii) of definition 1.1.1).

<u>Remark A-1</u>: It is sometimes convenient to be able to define the q-functions by prescribing their values only when the sets $A_1, A_2, \ldots, A_m$ are disjoint. Suppose, we have a set of functions $q_o(A_1, \ldots, A_m; r_1, \ldots, r_m)$ defined whenever the sets $A_1, \ldots, A_m$ are disjoint so that q_o can be regarded as defining the joint distribution of the random variables $\xi(A_1, \omega), \ldots, \xi(A_m, \omega)$ whenever the sets A_i's are disjoint. At this point it is not clear how condition A-1(iii) is defined.

Suppose that condition A-1(i) and (ii) are satisfied for disjoint A_i's and also that condition A-1(iv) is satisfied. Suppose further that if $A_1, A_2, \ldots, A_m$ are disjoint sets, each being a union of a finite number of disjoint sets i.e., $A_i = A_{i1} \cup A_{i2} \cup \ldots$ then the joint distrubition of $\xi(A_1, \cdot), \ldots, \xi(A_m, \cdot)$ is the same as that of

$$\sum_j \xi(A_{1j}, \cdot), \ldots, \sum_j \xi(A_{mj}, \cdot).$$

For example, if A, B and C are disjoint sets then we require

$$\text{(A-1-1)} \quad q_o(A, B \cup C;\ r_1, r_2) = \sum_{r_3 + r_4 = r_2} q_o(A, B, C;\ r_1, r_3, r_4).$$

With this definition the functions q_o can be extended in a unique manner to functions q that agree with q_o when the sets A_i are disjoint. For completion of the proof, interested readers are referred to Harris [1963].

<u>Theorem A-1</u>: Let $q(A_1, \ldots, A_m; r_1, \ldots, r_m)$ be given defined whenever $A_1, \ldots, A_m \in \mathcal{F}_{\mathcal{D}}$ and satisfying conditions A-1(i-iv) when the sets involved are in $\mathcal{F}_{\mathcal{D}}$. Then there exists a unique probability measure Q on C^* such that

$Q\{\xi(A_1,\omega) = r_1,\ldots,\xi(A_m,\omega) = r_m\} = q(A_1,\ldots,A_m,r_1,\ldots,r_m), r_1,\ldots,r_m = 0,1,\ldots$

whenever the A's $\in F_{\mathcal{D}}$.

Proof: The fundamental theorem of Kolmogorov [1956, p. 29] implies that the q-functions determine a unique probability measure Q_1 on $C^*(\Omega')$.

Claim: $\Omega'' \in C^*(\Omega')$ and $Q_1(\Omega'') = 1$

Now Ω'' consists of those $\xi(A,\omega)$ of Ω' that satisfy

$$\xi\left(\bigcup_{i=1}^{m} A_i,\omega\right) = \sum_{1}^{m} \xi(A_i,\omega)$$

where $A_1,\ldots,A_m \in F_{\mathcal{D}}$ and are disjoint. By condition A-1(iii) each such relation has Q_1-measure 1 and there are only denumerably many of them.

Thus it follows that

$$\Omega'' \in C^*(\Omega') \text{ and } Q_1(\Omega'') = 1.$$

Further, now a cylinder set B in Ω'' is the intersection of Ω'' with a cylinder set in Ω' so that

$$B = B_1 \cap \Omega'' \text{ where } B_1 \in C^*(\Omega')$$

Thus, having a probability measure Q_1 on $C^*(\Omega')$, we can define a probability measure Q' on $C^*(\Omega'')$ by putting: $Q'(B) = Q_1(B_1)$ for $B \in C^*(\Omega'')$. Conversely if Q' is a probability measure on $C^*(\Omega'')$ then we can define a probability measure Q_1 on $C^*(\Omega')$ by putting

$$Q_1(B_1) = Q'(B_1 \cap \Omega'') \quad \text{for } B_1 \in C^*(\Omega')$$

Also, the measure Q' is unique since if not, there would be two different Q_1 measures on $C^*(\Omega')$ contradicting the uniqueness of Kolmogorov theorem.

Thus, we have a unique probability measure Q' on $C^*(\Omega'')$ and $Q'(\Omega'') = 1$. If now $\xi \in \Omega''$ is such that it does not satisfy b(ii) for $A_i \in F_D$ namely

$$A_1 \supset A_2 \supset \ldots \quad \text{with} \quad \cap A_n = \phi$$

$$\xi(A_n,\omega) \not\to 0 \quad \text{for} \quad A_1,\ldots \in F_D$$

then the set Ω_o of all such ξ's has a measurable subset whose Q_1 measure is zero (follows from condition A-1(iv). This implies that,

$$Q_{1^*}(\Omega_o) = \text{inner measure of } \Omega_o = 0$$

clearly we have $\Omega''' = \Omega'' - \Omega_o$ so that $Q_1^*(\Omega''') = 1$.
(Q_1^* denotes the outer probability measure)

As discussed before, having a unique probability measure on $C^*(\Omega'')$ we can have a unique probability measure Q on $C^*(\Omega''')$ by putting $Q(B \cap \Omega''') = Q'(B)$: $B \in C^*(\Omega'')$. Consequently it follows that there exists a unique probability measure Q determined by the given q-functions.

<u>Remark A-1</u>: The easy implication of theorem 1.1.1 says that every Q on $(M_+(S), M(S))$ can be regarded as a P_ξ for some point process ξ, and the converse implication as proved here shows that given a point process ξ there exists a unique probability measure Q determined by the finite dimensional distributions of $\xi(A,\omega)$ given by the q-functions.

APPENDIX B

Theorem B-1: A sequence $\{X_i: i \geq 1\}$ of independent identically distributied random variables is strictly stationary and ergodic.

Proof: a) To show that $\{X_i: i \geq 1\}$ is stationary we need to show that $(X_1,X_2,\ldots)$ has the same joint distribution as $(X_2,X_3,\ldots)$ This follows clearly from the identically distributed property of the X_i's.

b) Need to establish the ergodicity of the stationary sequence $\{X_i: i \geq 1\}$. For each $\omega = (x_1,x_2,\ldots,x_n,\ldots)$ define a new process by

$Y_i(\omega) = X_i$ the ith-coordinate the $\{Y_i: i \geq 1\}$ defines a stochastic process known as the co-ordinate representation of $\{X_i: i \geq 1\}$ clearly both $\{X_i: i \geq 1\}$ and $\{Y_i: i \geq 1\}$ have the same distribution. Define a transformation S by

$S(x_1,x_2,\ldots,x_n,\ldots) = (x_2,x_3,\ldots,x_{n+1},\ldots)$ (S is the so-called shift transformation).

In what follows have a probability space $(\mathcal{R}_\infty,C_\infty,\hat{P})$ where $\hat{P}$ is defined by

$$\hat{P}(B) = P[(x_1,x_2,\ldots) \in B].$$

(i) S is measurable for:

Let C be a measurable finite dimensional product cylinder then $S^{-1}C$ is also such a cylinder set. These cylinder sets generate C_∞ implies that S is $\mathcal{C}_\infty$-measurable

(ii) $$\hat{P}(S^{-1}C) = \hat{P}(C) \quad \forall\ C \in \mathcal{C}_\infty$$

For: Let $C \in \mathcal{C}_\infty$.

$$\begin{aligned}
\hat{P}(S^{-1}C) &= \hat{P}[\omega \mid S\omega \in C] \\
&= \hat{P}[x_1,x_2,\ldots) \mid (x_2,x_3,\ldots) \in C] \\
&= P[(x_2,x_3,\ldots) \in C] \\
&= P[(x_1,x_2,\ldots) \in C] \quad \text{(by stat.)} \\
&= \hat{P}(C)
\end{aligned}$$

(i) and (ii) together imply that S is measure-preseving.

(iii) Need to show that $P(A) = 0$ or 1 for invariant events A.

Let A be an invariant event and let $\mathcal{G}_n = (X_m\colon\ m \geq n)$.

Let $\mathcal{G} = \bigcap_n \mathcal{G}_n$.

A invariant and S measure-preserving implies that $S^{-1}A = A$

implies $$[\omega \mid (x_2,x_3,\ldots) \in A] = [\omega \mid (x_1,x_2,\ldots) \in A].$$

Continuing this gives:

$$[\omega \mid (x_k,x_{k+1},\ldots) \in A] = [\omega \mid (x_1,x_2,\ldots) \in A] \ \forall\ k \geq 1$$

implies $A \in \mathcal{G}_k$ $\forall\ k \geq 1$ implies $A \in \mathcal{G}$, the tail σ-field. Since the X_i's are independent, by Kolmogorov 0-1 law it follows that

$$P(A) = 0 \ \text{ or } \ 1$$

(i), (ii), and (iii) imply (b)

Theorem B-2: [Anderson - 1971, p. 427]

Let $y_1, y_2, \ldots$ be a stationary stochastic process such that for every integer n and integers $t_1, \ldots, t_n$ $(0 < t_1 < \ldots < t_n)$ $y_{t_1}, \ldots, y_{t_n}$ is distributed independently of $y_1, \ldots, y_{t_1-m-1}$ and $y_{t_n+m+1}, \ldots$. If $Ey_t = 0$ and $Ey_t^2 < \infty$ then $\frac{1}{\sqrt{T}} \sum_1^T y_t$ has a limiting normal distribution with mean zero and variance

$$Ey_1^2 + 2E\, y_1 y_2 + \ldots. + 2E\, y_1 y_{m+1}$$

Theorem B-3: [Hannan - 1970, p. 203]

If $\{x(n)\}$ is stationary and ergodic and $E\{|x_j(n)|\} < \infty$

then $$\lim_{N\to\infty} \frac{1}{N} \sum_1^N x(n) = E\{x(n)\} \text{ a.s.}$$

Also if $$E\{(x_j(n)^2\} < \infty$$

then

$$\lim_{N\to\infty} \frac{1}{N} \sum_1^N x(m)x(m+\ell) = E\{x(m)x(m+\ell)\}$$

Remark B-1: Proof of the fact that

$$\epsilon_n = \frac{1}{n} [L_n''(0) - L_n''(\beta^*)] \text{ converges}$$

a.s. to zero as $n \to \infty$.

Proof: Let $\epsilon > 0$ be given.

Let the parameter space $\mathcal{B}$ be an open interval of the real line. Assume that $\forall\ \beta \in \mathcal{B}$ there corresponds an $\eta(\beta) > 0$ such that

$$(*) \qquad E_\beta[\sup_\beta\{|\frac{\partial^2 h(X_0,X_1,X_2;\ \beta^*)}{\partial\beta^2}|:\ \beta^* \in \mathcal{B},\ |\beta-\beta^*| < \eta(\beta)\}]$$

is finite where E_β denotes the expectation when β is the true parameter.

Let $U_0 = \{\beta:\ |\beta| \leq \eta_0 = \eta(0)\}$ be a neighbourhood of $\beta = 0$

Choose a δ: $0 < \delta < \eta_0$.

Now

$$h(X_0,X_1,X_2;\ \beta) = [\alpha + \beta(X_0 + X_2)]\ X_1 - \ln\,[1 + e^{\alpha+\beta(X_0 + X_2)}]$$

clearly $\frac{\partial h(X_0,X_1,X_2;\ \beta)}{\partial\beta}$ and $\frac{\partial^2 h(X_0,X_1,X_2;\ \beta)}{\partial\beta^2}$

exist and are continuous

Further, the continuity of $\frac{\partial^2 h(\ ;\beta)}{\partial\beta^2}$ implies the lower semi-continuity of

$$\sup\{|\frac{\partial^2 h(X_0,X_1,X_2;\ \beta)}{\partial\beta^2} - \frac{\partial^2 h(X_0,X_1,X_2;\ 0)}{\partial\beta^2}|:\ |\beta| < \delta\}$$

From assumption (*) we have for $\delta > \eta_0$

$$E_0\{\sup|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}|:\ |\beta^*| < \delta\} < \frac{\epsilon}{2}$$

Let us choose such a δ. Now under H_0, $\hat{\beta}_n \xrightarrow{a.s.} 0$ so that there exists an integer N_1 such that $\forall\ n \geq N_1$ $|\hat{\beta}_n| < \delta$

(N_1 possibly depending on the sample).

Thus for $n \geq N_1$:

$$|\epsilon_n| = n^{-1}|L_n''(0) - L_n''(\beta^*)|$$

$$\leq n^{-1} \sum_1^n \left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right|$$

$$\leq n^{-1} \sum_1^n \sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < |\hat{\beta}_n|\}$$

$$\leq n^{-1} \sum_1^n \sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < \delta\}.$$

By assumption (*)

$$E_0\{\sup\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < \delta\}$$

is finite.

Since under H_0 the sequence $\{X_i : i \geq 1\}$ is stationary and ergodic theorem 2.1[2] of Borwanker et al implies that

$$n^{-1} \sum_1^n \sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < \delta\}$$

converges a.s. to

$$E_0[\sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta)}{\partial\beta^2}\right| : |\beta| < \delta\}$$

Therefore there exists an integer N_2 such that for $n > N_2$

$$n^{-1} \sum_1^n \sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < \delta\}$$

$$< 2\, E_0\{\sup\{\left|\frac{\partial^2 h(\ ;\ 0)}{\partial\beta^2} - \frac{\partial^2 h(\ ;\ \beta^*)}{\partial\beta^2}\right| : |\beta^*| < \delta\}\}$$

N_2 again possibly depending on the sample

Take $N = \max(N_1, N_2)$ Thus for $n > N$

$$|\epsilon_n| < \epsilon$$

ϵ being arbitrary, it follows that $|\epsilon_n| \to 0$ a.s.

APPENDIX C

The conditional probability distribution of a 1-dimensional near-neighbour auto-binary spatial secheme is given by

$$(C\text{-}1) \qquad P_\beta[X_i|N\text{-}N] = \frac{e^{[\alpha+\beta(X_{i-1}+X_{i+1})]X_i}}{1+e^{\alpha+\beta(X_{i-1}+X_{i+1})}}$$

where β describes the spatial interaction between near-neighbour particles.

The hypotheses of interest as given by (4.2.1) are

$$\begin{cases} H_0: & \beta = 0 \\ H_{a,n}: & \beta = \beta_n = n^{-\frac{1}{2}} \to 0 \quad \text{as} \quad n \to \infty. \end{cases}$$

We need to show that in this formulation the LAN conditions given by (4.1.7) are not satisfied.

Now

$$\Lambda_n(\beta_n,0) = L_n(\beta_n) - L_n(0)$$

Using Taylor's series expansion we have

$$\Lambda_n(\beta_n,0) = \beta_n\, L_n'(0) + \frac{\beta_n^2}{2} L_n''(0) + \frac{\beta_n^3}{6} L_n'''(\beta^*)$$

where $|\beta^*| < |\beta_n| = n^{-\frac{1}{2}}$.

Following the argument of Section 3.2, we get

$$\Lambda_n(\beta_n,0) = \frac{1}{\sqrt{n}} \{\sum_{i=1}^{n} X_i(X_{i-1} + X_{i+1}) - \frac{e^{\alpha}(X_{i-1} + X_{i+1})}{1 + e^{\alpha}}\}$$

$$- \frac{1}{2n} \sum_{1}^{n} \frac{(X_{i-1} + X_{i+1})^2 e^{\alpha}}{(1 + e^{\alpha})^2} + \frac{n^{-\frac{1}{2}}}{6} \{\frac{1}{n} L_n'''(\beta^*)\}$$

By claim 3 (of Section 3.2) $\frac{1}{n} L_n'''(\beta^*)$, with $|\beta^*| < |\beta_n| = n^{-\frac{1}{2}}$, is asymptotically bounded (under H_0) as $n \to \infty$.

By claim 2 (of Section 3.2) the 2nd term on the right $\xrightarrow{a.s.}$ $-k^2$ (under H_0)

Further,

$$T_n(0) = \frac{1}{\sqrt{n}} \sum_{1}^{n} \{X_i(X_{i-1} + X_{i+1}) - \frac{e^{\alpha}(X_{i-1} + X_{i+1})}{1 + e^{\alpha}}\}$$

is asymptotically normal with mean zero and variance $\sigma_\xi^2 = \frac{e^{2\alpha}(3 + e^{\alpha})}{(1 + e^{\alpha})^4}$ (under H_0).

Thus, clearly (4.1.7) (b) is satisfied with

$$\Gamma_0 = \frac{e^{2\alpha}(3 + e^{\alpha})}{(1 + e^{\alpha})^4}$$

However (4.1.7) (a) is not satisfied for

$$\Lambda_n(\beta_n,0) - T_n(0) + \tfrac{1}{2}\Gamma_0 \xrightarrow{P_{n,0}} -k^2 + \tfrac{1}{2}\sigma_\xi^2$$

$$\ell.h.s. = -k^2 + \tfrac{1}{2}\sigma_\xi^2 = \frac{-e^{2\alpha}(1 + 2e^{\alpha})}{(1 + e^{\alpha})^4} + \tfrac{1}{2}\frac{e^{2\alpha}(3 + e^{\alpha})}{(1 + e^{\alpha})^4}$$

$$= \frac{e^{2\alpha}(1 - 3e^{\alpha})}{(1 + e^{\alpha})^4}$$

$$\neq 0 \text{ for all } \alpha$$

Thus,

$$\Lambda_n(\beta_n,0) - T_n(0) + \tfrac{1}{2}\,\Gamma_0 \overset{P_{n,0}}{\not\longrightarrow} 0 \ \forall\ \alpha$$

Consequently, it follows that LAN conditions are not satisfied in our 1-dimensional auto-binary spatial scheme. In a similar manner, it can be shown that the LAN conditions are not satisfied in our 2-dimensional isotropic model either. Thus, the classical techniques of contiguity fail in our formulation.

REFERENCES

1. Anderson, T.W.: (1971)
   The statistical analysis of time series, John Wiley, N.Y.

2. Berge, Claude: (1962)
   The theory of graphs and its applications, John Wiley, N.Y.

3. Besag, Julian: (1974)
   Spatial interaction and the statistical analysis of lattice systems (with discussion), J.R.S.S. B, 36, 192-236.

4. Besag, J. and Gleaves: (1973)
   On the detection of spatial pattern in plant communities, Bull. Inst. Stat. Ins. 45(1), 153-8.

5. Billingsley, P: (1968)
   Convergence of probability measures, John Wiley, N.Y.

6. Borwarker, Kallianpur, Prakasa Pao: (1971)
   The Bernstein-Von Mises theorem for Markov processes, Ann. Stat 42, 1242-1253.

7. Carnal, Etienne: (1979)
   Processes markoviens a plusieurs parameters, dissertation.

8. Cramer, Harold: (1946)
   Mathematical methods of statistics, Princeton University Press.

9. Diggle, Besag and Gleaves: (1976)
   Statistical analysis of spatial patterns by means of distance methods, Biometrics 32, 659-667.

10. Friedman:
    Stochastic differential equations and applications, Academic Press.

11. Grimmett, G.R.: (1973)
    A theorem about random fields, Bull London Math. Soc. 5, 81-84.

12 Hajek and Sidek: (1967)
    The theory of rank tests, Academic Press.

13. Hannan, E.J.: (1970)
    Multiple time series, John Wiley, N.Y.

14. Harris, E.T.: (1963)
The theory of branching processes, Springer-Verlag.

15. Holgate, P: (1965)
Test of randomness based on distance methods, Biometrika, 52, 345-53.

16. Kolmogorov, A.: (1956)
Foundations of the theory of probability, N.Y. Chelsea Publishing Company.

17. Matern, B.: (1960)
Spatial Variation.

18. Neveu, J.: (1976)
Processes Pontuels, Lecture notes in Mathematics #598, Springer-Verlag.

19. Pettis, Bailey, Jain and Dubes: (1979)
An intrinsic dimensionality estimator from near-neighbour information, IEEE transactions on pattern analysis and machine intelligence, Vol. PAM I-1, #1, 25-36.

20. Pielou, E.C.: (1959)
The use of point-to-plant distances in the study of the pattern of plant populations, J. Ecology, 47, 607-613.

21. Preston, C.J.: (1974)
Gibbs states on countable sets, Cambridge University Press, London.

22. Roussas, G.G.: (1972)
Contiguity of probability measures: Some applications in statistics, Cambridge Press.

23. Spitzer, F.: (1971)
Markov random fields and Gibbs ensembles, Amer. Math Mthly. 78, 142-154.

24. Srivastava, J.N.: (1976)
An information function approach to dimensionality analysis and curved manifold clustering, Mult. Analy-III, P.R. Krishnaiah, 369-382.