総合都市研究　特別号　1989

点分布パターン分析技法とその FORTRAN プログラム

杉　浦　芳　夫*

要　　約

本稿では，空間的データの解析に有用な点分布パターン分析技法（最近接尺度、空間的連関係数，ユークリッド２次元回帰分析）の説明を行なった。そして，それらの FORTRAN プログラムを添付するとともに，簡単な事例の出力結果も示した。

1．はじめに

1950－1960年代の地理学における計量革命が斯学にもたらした帰結の一つは，空間的分布パターンの正確な把握への関心の高まりとその分析技法の開発であった（杉浦，1989)。この結果，空間的分布パターンを，点，線，面といった幾何学的構成要素から定量的に測定することが可能となった。本稿では，特に点分布パターン分析技法のうち，最近接尺度 Nearest neighbor method（King, 1962)，空間的連関係数 Coefficient of spatial association（Sorensen, 1974)，ユークリッド２次元回帰分析 Euclidean bidimensional regression analysis（Tobler, 1965）の三つについての説明を行なうとともに，それらの FORTRAN プログラムを示すことにしたい。

2．最近接尺度

離散的な点の分布パターン（例えば都市や工場の分布パターン）の典型タイプとしては，凝集分布，ラムダム（無秩序）分布，均等分布の三つがある（図１)。ここで問題とすることは，実際の点分布パターンが，これらの三つの典型的分布パターンのいずれと最もよく類似しているかを客観的に判定することである。そのための方法としては，一つは，対象地域全体にメッシュ（方格）をかぶせて，各メッシュに含まれる点の数を数え，その度数分布を特定の確率分布（例えばポアソン分布）で近似させ，点分布パターンの判定を行なう方格法 Quadrat method（Rogers, 1974）がある。他方，最近接尺度は，各点の最近接に位置する点までの距離を測定することによって，点分布パターンを判定しようとするものである。問題とする点分布が，ラムダム分布と有意な差があるか否かをパターン判定の目安にしている点では両者は共通している。ただし，前者の方法が便宜的に各メッシュに含まれる点の数を数えるのに対し，後者の方法は，点間距離を直接測定するという違いがある。

＊東京都立大学都市研究センター・理学部

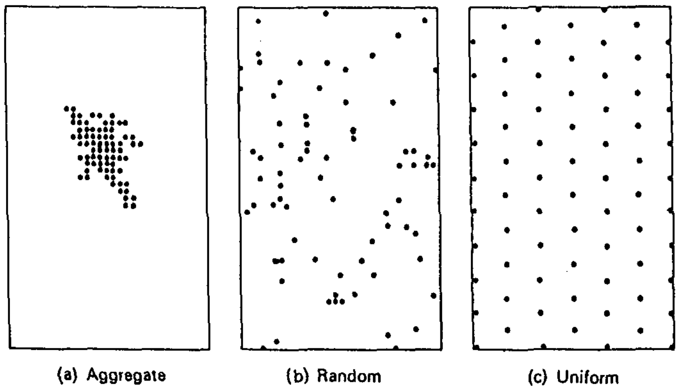

図１　空間的分布パターンの典型例
出典：King（1962）

このうち，方格法は，メッシュの大きさを変えることによって判定結果が異なる可能性があるのに加え，点間の空間的位置関係に全く考慮が払われていない欠点がある。空間的分布という場合の分布 Distribution は，統計学的な度数分布とは異なり，空間的な関係を含意している。すなわち，点の配列 Arrangement こそが問題とされているにもかかわらず，方格法は，配列に関して敏感ではないのである。それに対し，最近接尺度はこれらの欠点から解き放たれている。最近接尺度は，元々は生態学者の Clark and Evans（1954）によって植生群落や動物集団の分布を測定するために考案された。最近接尺度の定義式は次のとおりである。

$$R = r_A / r_E \tag{1}$$

ただし，r_A は，現実に分布する n 個の点に関し，各点の最近接点までの直線距離 r_i の平均値，すなわち

$$r_A = \sum_{i=1}^{n} r_i / n \tag{2}$$

である。r_E は，理論的なラムダム分布（ポアソン分布）を前提とした場合の，n 個の点に関し，各点の最近接点までの直線距離の平均値であり，次式で定義される（詳しい証明は Clark and Evans（1954），奥野（1977）を参照されたい）。

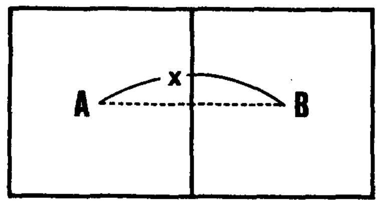

図２　正四角形パターン

$$r_E = 1/2\sqrt{n/A} \tag{3}$$

ただし，A は対象地域の面積であるため，n/A は点の密度に相当する。

もし現実の点分布が完全なラムダム・パターンであるなら，r_A と r_E が一致するため，第(1)式より，$R=1$ となる。また，すべての点が一点に集中する完全な凝集パターンの場合には，$r_A=0$ であり，第(1)式より $R=0$ となる。他方，均等分布の代表としては，点が正四角形のメッシュの中心に分布する正四角形パターンと，点が正六角形のメッシュの中心に分布する正六角形パターンがあり，次のようにしてそれぞれの最近接尺度値が推定される。

まず正四角形パターン（図２）の場合は，点Ａ・Ｂ間の距離を x とすれば，一つの正四角形の面積は x^2 である。そして，対象地域全体がこうした正四角形メッシュで完全におおいつくされていると，地域の面積 A を点の数 n で除したものが一つの正四角形の面積に等しい。すなわち，x^2

$=A/n$ である。したがって，この場合の点の最近接距離は，

$$r_A = x = \sqrt{A/n} \tag{4}$$

となる。そこで，第(1)式に第(3)・(4)式を代入すると，

$$R = r_A / r_E = \sqrt{A/n} / (1/2\sqrt{n/A}) = 2$$

となる。この値が，点が正四角形パターンの均等分布を呈する場合の最近接尺度値である。

次に，正六角形パターン（図3）の場合は，点A・B間の距離を x とすれば，$AC=x/2$，$CD=x/2\sqrt{3}$ であるため，一つの正六角形の面積は，$x/2 \times x/2\sqrt{3} \times 1/2 \times 12 = \sqrt{3}x^2/2$ となる。そして，対象地域全体がこうした正六角形メッシュで完全におおいつくされていると，地域の面積 A を点の数 n で除したものが一つの正六角形の面積に等しい。すなわち，$A/n=\sqrt{3}x^2/2$ である。したがって，この場合の点の最近接距離は，

$$r_A = x = \sqrt{(2/\sqrt{3}) \times (A/n)} \tag{5}$$

となる。そこで，第(1)式に第(3)・(5)式に代入すると，

$$R = r_A / r_E = \sqrt{(2/\sqrt{3}) \times (A/n)} / (1/2\sqrt{n/A})$$
$$= \sqrt{8/\sqrt{3}} = 2.149$$

となる。この値が，点が正六角形パターンの均等分布を呈する場合の最近接尺度値である。

以上の尺度値の比較から，最も代表的な均等分布が正六角形型のものであるとすると，現実の点分布は，最小0（完全凝集分布）から最大2.149（完全均等分布）までの最近接尺度値をとることになる。最近接尺度では，ラムダム・パターンとの比較において，問題とする点分布パターンの判定を行なうため，求められた尺度値と完全ラムダム分布（$R=1$）の場合のそれとの有意差の検定がなされねばならない。この検定には，平均0，分散1の標準正規分布を呈する Z スコアを用いる。

$$Z = (r_A - r_E) / \sigma_{r_E} \tag{6}$$

ただし，σ_{r_E} はポアソン分布から導出される標準誤差であり，次式で定義される（詳しい証明は Clark and Evans (1954)，奥野 (1977) を参照されたい)。

$$\sigma_{r_E} = 0.26136 / \sqrt{n \times (n/A)} \tag{7}$$

具体的なパターン判定は次のようにして行なう。所定のデータから求められた第(6)式の Z スコアと，一定の有意水準のもとでの正規確率限界値 Z_α（ただし，$Z_{.05}=1.960$，$Z_{.01}=2.576$ である）を比較し，$Z_\alpha > |Z|$ ならば帰無仮説（当該点分布はラムダム分布と有意差がない）を採択し，$Z_\alpha \leq |Z|$ ならば帰無仮説を棄却する。ただし，この統計量は，標準誤差がポアソン分布から導出されているため，ラムダム分布（$R=1$）との有意差は検定できても，凝集分布（$R=0$）や均等分布（$R=2.149$）との有意差の検定はなしえない。したがって，例えば $R=1.8$ という値が求められ，当該分布とラムダム分布に有意差のあることが判明したとしても，それがすぐに均等分布と判定されることにはならず，単に点分布が均等分布に近いことを示唆しているにすぎないのである。

最近接尺度の FORTRAN プログラムと出力例は付図1に示すとおりである。入力データとしては，点の総数，対象地域の面積をまず最初に用意し，次に，各点の $x \cdot y$ 座標を用意すればよい。出力は，入力データに続いて，現実の点間平均距

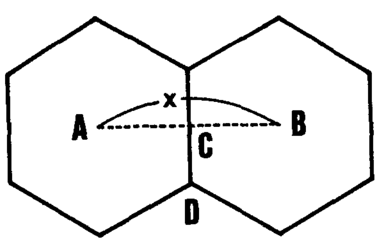

図3　正六角形パターン

離，理論上の点間平均距離，最近接尺度値，標準誤差，Zスコア，ラムダム分布との有意差検定結果が印刷される。

3．空間的連関係数

最近接尺度は，当該点分布が三つの典型的な点分布のいずれに近似しているかについてパターン判定を行なうものであった。それとは別に，二つの異なる点分布の空間的パターンの類似度を測定することが要請される場合もある。例えば，実際の点分布とシミュレートされた点分布の空間的パターンの類似度の測定は，空間的モデルの検証に不可欠である（杉浦，1975)。このような目的のために開発されたものが，Sorensen（1974）の空間的連関係数である。

空間的連関係数も基本的には最近接距離を測定することによって，次のようにして求められる。いま二つの点分布A，Bがあるとすると，1）点分布Aにおける各点の他の点までの最近接距離をdA_i，2）点分布Bにおける各点の他の点までの最近接距離をdB_i，3）点分布Aにおける各点の点分布Bにおける点までの最近接距離をdAB_i，4）点分布Bにおける各点の点分布Aにおける点までの最近接距離をdBA_iとし（図4），二つの点分布における各点分布内での平均最近接距離を第(8)式，ならびに一方の点分布における各点の他方の点分布の点までの平均最近接距離を第(9)式で定義すれば，

$$(\sum_{i=1}^{n} dA_i + \sum_{i=1}^{m} dB_i)/(n+m) = a \qquad (8)$$

$$(\sum_{i=1}^{n} dAB_i + \sum_{i=1}^{m} dBA_i)/(n+m) = b \qquad (9)$$

基準化されていない空間的連関係数C'_sは，両者の比をとって，第(10)式で定義される。

$$C'_s = a/b = (\sum_{i=1}^{n} dA_i + \sum_{i=1}^{m} dB_i)/(\sum_{i=1}^{n} dAB_i + \sum_{i=1}^{m} dBA_i) \qquad (10)$$

ただし，nは点分布Aの点の総数，mは点分布Bの点の総数である。

C'_sは0から∞までの値をとり，値が大きくなる程，二つの点分布パターンの類似度が増大することが経験的に確かめられている。しかし，C'_sの値は点の総数に依存しているため，何等かの基準化が必要である。そこで，点分布パターンの相互比較を可能にする意味でも，C'_sが相関係数のように－1から＋1までの値をとる範囲で変動するような変換を考える。経験的には，C'_sは類似度の増大に伴って指数関数的に増加することが判明している（Sorensen, 1974)。指数関数的な曲線で近似される値の変動幅を絶対値1以下にするために，ここでは次のような変換を施して基準化し，それを空間的連関係数C_sとよぶことにする。

$$C_s = (a-b)/(a+b) = \{(\sum_{i=1}^{n} dA_i + \sum_{i=1}^{m} dB_i) - (\sum_{i=1}^{n} dAB_i + \sum_{i=1}^{m} dBA_i)\}/\{(\sum_{i=1}^{n} dA_i + \sum_{i=1}^{m} dB_i) + (\sum_{i=1}^{n} dAB_i + \sum_{i=1}^{m} dBA_i)\} \qquad (11)$$

このような基準化を行なうと，C_sは－1と＋1の間の値をとる。そして，そのパターン判定のおおよそ基準は次のようである。

$C_s < -0.5$　強い非類似性
$-0.5 \leq C_s < -0.2$　ある程度の非類似性

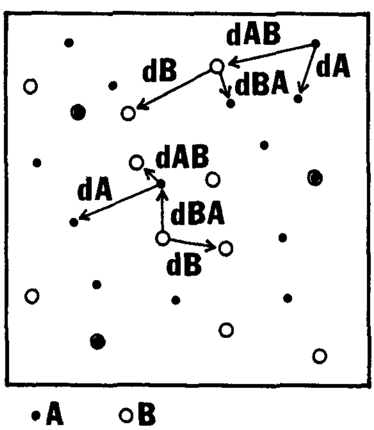

図4　分布内最近接距離と分布間最近接距離

$-0.2 \leqq C_s < +0.2$　顕著な類似性・非類似性なし
$+0.2 \leqq C_s < +0.5$　ある程度の類似性
$+0.5 \leqq C_s$　強い類似性

空間的連関係数の特徴は，二つの点分布の点が必ずしも同数である必要がないことである。ただし，えられた係数値の有意性を判定する統計量がないことが大きな欠点である。

空間的連関係数の FORTRAN プログラムと出力例は付図２に示すとおりである。入力データとしては，一方の点分布における点の総数と他方のそれをまず最初に用意し（ただし，二つの点分布の点の総数が同じなら，二番目の点の総数は省略する），次に，まず最初の点分布の各点の $x \cdot y$ 座標，続いて二番目の点分布の各点の $x \cdot y$ 座標をそれぞれ別々に用意すればよい。出力は，入力データと空間的連関係数値が印刷される。

4．ユークリッド２次元回帰分析

二つの点分布を比較する際，一方の点分布の各点を他方の点分布のいずれか一つの点に対応づけてパターンの類似度を判定することが必要な場合がある。例えば，現実の都市の立地地点と理論上のそれを対応づけ，どの程度まで現実の都市の立地パターンが理論上のそれに類似しているかを検討する場合である（Evans and Gould, 1982)。これは，いわば一方の地図をもう一方の地図へ変換する地図変換の問題でもある（杉浦，1988)。

二つの点分布の類似度は，一方の点分布を他方の点分布の関数とみなすことによって可能となる。すなわち，点分布Wの座標を (u, v)，点分布Zの座標を (x, y) とすると，

$$\left.\begin{aligned} u &= \mathrm{f}(x, y) \\ v &= \mathrm{g}(x, y) \end{aligned}\right\} \quad (12)$$

なる二つの関数を求め，これを用いた座標変換によって点分布Zが点分布Wを近似する度合が大きいほど，二つの点分布は類似度が大きいとみなすのである。関数としては様々なものが考えられる

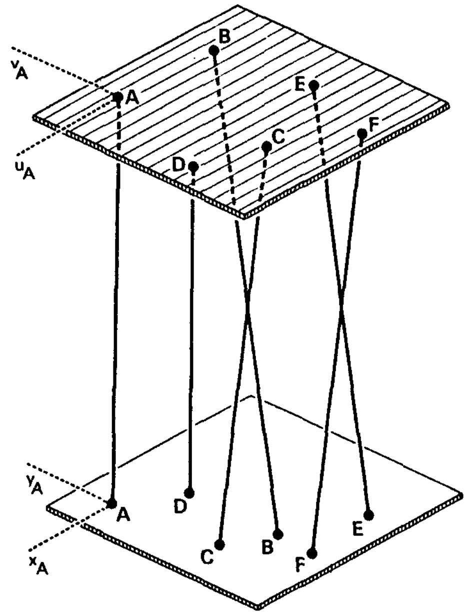

図５　ユークリッド２次元回帰の考え方
出典：Gatrell (1983)

が，最も単純な場合である線形関数を用いて，２次元平面上の点分布Zと点分布Wの類似度ないしは相関を測定するとともに，前者によって後者を可能な限り正確に予測しようとするものが，Tobler（1965）のユークリッド２次元回帰分析である（図５)。

いま，座標平面として複素平面を考えると，複素数 $z = x + iy$ は，(x, y) で定まるため，z を (x, y) の点で対応させると，z はこの点の位置ベクトルとみなすことができる。以下では，複素平面を前提として，ユークリッド座標変換を可能とする関数式を誘導してみよう。複素平面上の点分布Wと点分布Zにおける点 j の位置は，

$$w_j = u_j + iv_j \quad (13)$$

$$z_j = x_j + iy_j \quad (14)$$

となる。誤差が存在することなく，線形関数によって点分布Zが点分布Wを完全に近似できるならば，

$$w_j = A + Bz_j \tag{15}$$

である。ただし，パラメータ A，B は，複素平面を前提としているため，

$$\begin{aligned} A &= a_1 + ia_2 \\ B &= b_1 + ib_2 \end{aligned} \tag{16}$$

である。

そこで，第(14)・(16)式を第(15)式に代入すると，

$$w_j = (a_1 + ia_2) + (b_1 + ib_2)(x_j + iy_j) \tag{17}$$

となる。さらに第(13)式を第(17)式に代入し，式を整理すると，

$$u_j + iv_j = a_1 + b_1x_j - b_2y_j + i(a_2 + b_2x_j + b_1y_j) \tag{18}$$

となる。第(18)式を実数部分と虚数部分にわけると，

$$\left.\begin{aligned} u_j &= a_1 + b_1x_j - b_2y_j \\ v_j &= a_1 + b_2x_j + b_1y_j \end{aligned}\right\} \tag{19}$$

となる。ここで，誤差 e_j，f_j を考慮して第(19)式を行列表示すると，

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{pmatrix} \begin{pmatrix} x_j \\ y_j \end{pmatrix} + \begin{pmatrix} e_j \\ f_j \end{pmatrix} \tag{20}$$

となる。これがユークリッド2次元回帰式であり，通常の単回帰式に類似した形を呈していることがわかる。

誤差を最小にして，パラメータ a_1，a_2，b_1，b_2 を推定するには，誤差の二乗和を最小にする最小二乗法を適用すればよい。この場合の，最小化すべき値は次のように定義される（Tobler, 1983）。

$$Q = \sum_j (w_j - \hat{w}_j)(w_j - \hat{w}_j)^* \tag{21}$$

ただし，$\hat{w}_j$ は第(17)式で推定される理論上の w_j である。また，$(w_j - \hat{w}_j)^*$ は $(w_j - \hat{w}_j)$ の共役複素数（$z = x + iy$ ならば $z^* = x - iy$）である。w_j に第(13)式を，$\hat{w}_j$ に第(17)式を代入すると，

$$\begin{aligned} w_j - \hat{w}_j &= (u_j + iv_j) - \{(a_1 + ia_2) + \\ &\qquad (b_1 + ib_2)(x_j + iy_j)\} \\ &= (u_j - a_1 - b_1x_j + b_2y_j) + \\ &\qquad i(v_j - a_2 - b_1y_j - b_2x_j) \end{aligned} \tag{22}$$

となる。ところで，

$$zz^* = (x + iy)(x - iy) = x^2 + y^2 \tag{23}$$

であるため，第(22)式より，第(21)式は，結局，

$$\begin{aligned} Q = \sum_j \{ &(u_j - a_1 - b_1x_j + b_2y_j)^2 + \\ &(v_j - a_2 - b_1y_j - b_2x_j)^2 \} \end{aligned} \tag{24}$$

となる。

Q の最小値を求めるには，a_1，a_2，b_1，b_2 を変数とみなし，Q を各々で偏微分し，それを0とおけばよい。その際に求められる a_1，a_2，b_1，b_2 の値が，Q に最小値を与えるべく推定されたパラメータ値である。そこで，Q を四つの変数でそれぞれ偏微分し，それを0とおくと次のようになる。

$$\begin{aligned} \frac{Q}{\partial a_1} &= -\sum_j (u_j - a_1 - b_1x_j + b_2y_j) = 0 \\ &na_1 + b_1\sum_j x_j - b_2\sum_j y_j = \sum_j u_j \end{aligned} \tag{25}$$

$$\begin{aligned} \frac{Q}{\partial a_2} &= -\sum_j (v_j - a_2 - b_1y_j - b_2x_j) = 0 \\ &na_2 + b_1\sum_j y_j + b_2\sum_j x_j = \sum_j v_j \end{aligned} \tag{26}$$

$$\begin{aligned} \frac{Q}{\partial b_1} &= \sum_j \{-x_j(u_j - a_1 - b_1x_j + b_2y_j) \\ &\quad -y_j(v_j - a_2 - b_1y_j - b_2x_j)\} = 0 \\ &\quad a_1\sum_j x_j + a_2\sum_j y_j + b_1\sum_j (x_j^2 + y_j^2) \\ &= \sum_j (u_jx_j + v_jy_j) \end{aligned} \tag{27}$$

$$\begin{aligned} \frac{Q}{\partial b_2} &= \sum_j \{y_j(u_j - a_1 - b_1x_j + b_2y_j) \\ &\quad -x_j(v_j - a_2 - b_1y_j - b_2x_j)\} = 0 \\ &\quad -a_1\sum_j y_j + a_2\sum_j x_j + b_2\sum_j (x_j^2 + y_j^2) \\ &= \sum_j (v_jx_j - u_jy_j) \end{aligned} \tag{28}$$

したがって，第(25)－(28)式の連立方程式（正規方程式）をとけば，a_1，a_2，b_1，b_2 が求められる。なお，正規方程式を行列表記すると次のようになる。

$$\begin{pmatrix} n & 0 & \sum_j x_j & -\sum_j y_j \\ 0 & n & \sum_j y_j & \sum_j x_j \\ \sum_j x_j & \sum_j y_j & \sum_j (x_j^2+y_j^2) & 0 \\ \sum_j y_j & \sum_j x_j & 0 & \sum_j (x_j^2+y_j^2) \end{pmatrix}$$

$$\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} \sum_j u_j \\ \sum_j v_j \\ \sum_j (u_j x_j + v_j y_j) \\ \sum_j (v_j x_j - u_j y_j) \end{pmatrix} \qquad (29)$$

第(29)式における逆行列をとかずに，手計算で a_1，a_2，b_1，b_2 を求めてみよう。第(25)式ならび第(26)式の両辺をそれぞれ n で除すと，

$$a_1 = \bar{u} - b_1\bar{x} + b_2\bar{y} \qquad (30)$$

$$a_2 = \bar{v} - b_1\bar{y} - b_2\bar{x} \qquad (31)$$

となる。そこで，第(27)式の両辺を n で除して，第(30)・(31)式の a_1，a_2 を代入し，式を整理すると，

$$b_1 = \{(1/n)\sum_j (u_j x_j + v_j y_j) - \bar{u}\bar{x} - \bar{v}\bar{y}\} / \{(1/n)\sum_j (x_j^2 + y_j^2) - \bar{x}^2 - \bar{y}^2\} \qquad (32)$$

となる。また，第(28)式の両辺を n で除して，第(30)・(31)式の a_1，a_2 を代入し，式を整理すると，

$$b_2 = \{(1/n)\sum_j (v_j x_j - u_j y_j) + \bar{u}\bar{y} - \bar{v}\bar{x}\} / \{(1/n)\sum_j (x_j^2 + y_j^2) - \bar{x}^2 - \bar{y}^2\} \qquad (33)$$

となる。あらかじめ求めておいた，$\bar{x}$，$\bar{y}$，$\bar{u}$，$\bar{v}$，$\sum_j x_j^2$，$\sum_j y_j^2$，$\sum_j u_j x_j$，$\sum_j v_j y_j$，$\sum_j v_j x_j$，$\sum_j u_j y_j$ を第(32)・(33)に代入すれば b_1，b_2 が求められ，さらにそれを第(30)・(31)式に代入すれば a_1，a_2 が求められる。

なお，第(32)式は，x_j と u_j の共分散と y_j と v_j の共分散の和を，x_j の分散と y_j の分散の和で除したものの変形であるため，次のようにも表わされる。

$$b_1 = \{\sum_j (x_j - \bar{x})(u_j - \bar{u}) + \sum_j (y_j - \bar{y})(v_j - \bar{v})\} / \{\sum_j (x_j - \bar{x})^2 + \sum_j (y_j - \bar{y})^2\} \qquad (34)$$

また，第(33)式は，x_j と v_j の共分散と，y_j と u_j の共分散の差を，x_j の分散と y_j の分散の和で除したものの変形であるため，次のようにも表わされる。

$$b_2 = \{\sum_j (x_j - \bar{x})(v_j - \bar{v}) - \sum_j (y_j - \bar{y})(u_j - \bar{u})\} / \{\sum_j (x_j - \bar{x})^2 + \sum_j (y_j - \bar{y})^2\} \qquad (35)$$

第(34)式で定義される b_1 は，丁度，通常の単回帰式の回帰係数が x と y の共分散を x の分散で除したものに等しいことと対比される。なお，付図3に示す FORTRAN プログラムでは，第(34)・(35)式を用いて b_1，b_2 が計算されている。

a_1，a_2，b_1，b_2 の別の求め方としては，第(19)式の二つの式を次のようにして，同時に最小二乗法の適用対象とみなし，次式で定義される Q' を a_1，a_2，b_1，b_2 でそれぞれ偏微分し，それらを0とおくことによってとく方法もある。これは，イメージ的には，通常の単回帰分析のパラメータ推定において適用される最小二乗法を第(20)式に適用することと等価である。

$$Q' = \sum_j \{u_j - (a_1 + b_1 x_j - b_2 y_j)\}^2 + \sum_j \{v_j - (a_2 + b_2 x_j + b_1 y_j)\}^2 \qquad (36)$$

四つのパラメータのうち，a_1，a_2 は通常の単回帰式の定数項に相当するものであり，(x_j，y_j) を (u_j，v_j) に重ねあわす際に行なう移動の大きさを示すパラメータである。他方，b_1，b_2 は通常の

単回帰式の回帰係数に相当するものであり，(x_j, y_j) を (u_j, v_j) に重ねあわす際に行なう拡大・縮小，さらには回転の大きさを示すパラメータである。ここにおいて，$b_1=1$，$b_2=0$ ならば，(x_j, y_j) は平行移動するだけで (u_j, v_j) に完全に一致する。

求められた四つのパラメータ値を用いれば，点分布Ｚから推定される理論上の点分布Ｗの座標 ($\hat{u}_j$, $\hat{v}_j$) が計算される。そして，点分布Ｗが点布Ｚによって説明される割合は，通常の単相関係数を二乗した決定係数と同様に定義される。決定係数は，１と，独立変数の変動によって説明されない従属変数の変動の割合の差，すなわち，１と，残差の分散を y の分散で除したものの差に等しい（野上・杉浦，1986)。これと同様に考えれば，２次元回帰分析の説明率は，

$$R^2=1-[\{\sum_j(u_j-\hat{u}_j)^2+\sum_j(v_j-\hat{v}_j)^2\}/\{\sum_j(u_j-\bar{u})^2+\sum_j(v_j-\bar{v})^2\}] \quad (37)$$

と定義され（Gatrell, 1983)，$0\leqq R^2\leqq 1.0$である。さらに，点分布Ｚと点分布Ｗの対応度，すなわち２次元相関係数は，通常の単相関係数と同様に，説明率の平方根 R で定義される。ただし，R の正負は，第(20)式の回帰係数相当行列の行列式の値の符号の正負で判定される（Tobler, 1965)。

ユークリッド２次元回帰分析の FORTRAN プログラムと出力例は付図３に示すとおりである。入力データとしては，点の総数に続いて，x・y 座標，u・v座標を各点ごとに対応させて用意すればよい。出力は，x・y 座標，u・v 座標，$\hat{u}$・$\hat{v}$ 座標，回帰係数相当行列の行列式の値，２次元相関数，説明率，ユークリッド２次元回帰式パラメータ値が印刷される。

5．適用事例

最後に，以上で紹介した方法の具体的適用事例を示しておこう。事例は，アメリカの文化人類学者 G. W. Skinner が，中心地理論（杉浦，1989）を援用して，1940年代の中国における市場町の立地を考察した有名なモノグラフ（スキナー，1979）からとったものである。図６中の白丸は，四川省成都の北東35－90kmに位置する金堂・中江両県にまたがる，1,840km² (46km×40km) の範囲に分布する19の市場町の位置を示している。

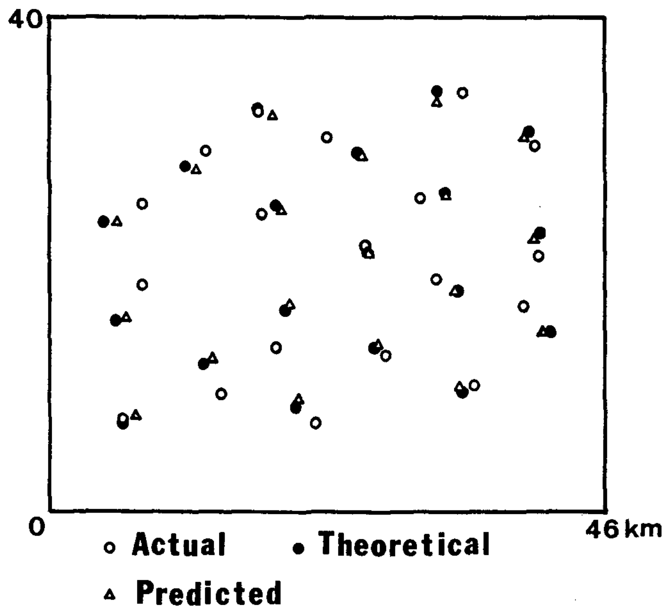

図６　中国・成都北東郊における市場町の分布

まず，この分布データに最近接尺度を適用してみると，尺度値が1.287，Zスコアが2.390（$>Z_{.05}=1.960$）となった。このことから，市場町の分布は5％水準でラムダム・パターンと有意な差があり，均等分布に近いことがわかる。

次に，中心地理論によって示唆されるところのものとしてSkinnerが示した，正六角形のメッシュの中心に立地する市場町の理論的位置（図6中の黒丸）と現実の位置との類似性を空間的連関係数で測定してみよう。係数値は0.575であり，両者の間に強い類似性があることがわかる。

最後に，市場町の理論上の位置（x_j，y_j）と現実の位置（u_j，v_j）との関係をユークリッド2次元回帰分析でとらえてみよう。2次元相関係数は0.988，理論上の位置による現実の位置の説明率は97.7％，回帰係数相当行列の行列式の値の符号は正であった。そして，ユークリッド2次元回帰式は，

$$\begin{pmatrix} u_j \\ v_j \end{pmatrix} = \begin{pmatrix} 0.994 \\ 1.051 \end{pmatrix} + \begin{pmatrix} 0.959 & 0.011 \\ -0.011 & 0.959 \end{pmatrix} \begin{pmatrix} x_j \\ y_j \end{pmatrix}$$

となった。なお，図6では，回帰式から予測された市場町の位置が三角で示されている。以上の結果から，中国四川省成都北東郊における市場町の分布は，中心地理論が示唆するような均等分布を呈しているといえよう。

文　献　一　覧

奥野隆史
1977　『計量地理学の基礎』大明堂.

杉浦芳夫
1975　「名古屋とその隣接地域における“アジアかぜ”の都市間拡散－空間的拡散研究の一事例として－」『地理学評論』48巻，pp. 847-867.

杉浦芳夫
1988　「分布図による表現－地図を歪める試み－」中村和郎・高橋伸夫（編）『地理学への招待』古今書院，pp. 47-61.

杉浦芳夫
1989　『立地と空間的行動』古今書院.

スキナー，G．W.
1979　『中国農村の市場・社会構造』（今井清一・中村哲夫・原田良雄訳）法律文化社.

野上道男・杉浦芳夫
1986　『パソコンによる数理地理学演習』古今書院.

Clark, Philip J. and Evans, Francis C.
1954 “Distance to nearest neighbor as a measure of spatial relationships in population.” *Ecology* 35：445-453.

Evans, Susan and Gould, Peter
1982 “Settlement models in archaeology.” *Journal of Anthropological Archaeology* 1: 275-304.

Gatrell, Anthony
1983 *Distance and Space: A Geographical Perspective.* Oxford: Clarendon Press.

King, Leslie J.
1962 “A quantitative expression of the pattern of urban settlements in selected areas of United States.” *Tijdschrift voor Econ. en Soc. Geographie* 53: 1 - 7 .

Rogers, Andrei
1974 *Statistical Analysis of Spatial Dispersion.* London: Pion.

Sorensen, Anthony D.
1974 “A method for measuring the spatial association between point patterns.” *Prof. Geogr.* 26: 172-176.

Tobler, Waldo R.
1965 “Computation of the correspondence of geographical patterns.” *Papers of the Regional Science Association* 15: 131-139.

Tobler, Waldo R.
1983 *Bidimensional Regression.* Discussion Paper 6 : Department of Geography, University of California, Santa Barbara.

Key Words（キー・ワード）

Point pattern（点分布パターン）, Nearest neighbor measure（最近接尺度）, Coefficient of spatial association（空間的連関係数）, Euclidean bidimensional regression analysis（ユークリッド２次元回帰分析）, FORTRAN program（フォートラン・プログラム）

付図１　最近接尺度の FORTARN プログラムと出力例

```
C
C *********************************************
C *     PROGRAM OF NEAREST NEIGHBOR MEASURE    *
C *                                            *
C *      (INPUT DATA)                          *
C *         N: NUMBER OF POINTS                *
C *         A: AREAL SIZE OF STUDY AREA        *
C *         X: X-COORDINATES                   *
C *         Y: Y-COORDINATES                   *
C *********************************************
C
      DIMENSION X(200),Y(200)
C
C     READ INPUT DATA
C
      READ(5,100) N,A
  100 FORMAT(I5,F10.2)
      WRITE(6,110) N,A
  110 FORMAT(' NUMBER OF POINTS = ',I5,'      AREA = ',F9.1,' KM**2 '//9
     1X,'X-COOR    Y-COOR'/)
      DO 1 I=1,N
      READ(5,120) X(I),Y(I)
    1 WRITE(6,130) I,X(I),Y(I)
  130 FORMAT(1X,I4,2F10.2)
  120 FORMAT(2F10.2)
C
C     CALCULATE THE OBSERVED MEAN DISTANCE
C
      FN=FLOAT(N)
      DMEAN=0.
      DO 2 I=1,N
      DMIN=9000000.0
      DO 3 J=1,N
      IF(I.EQ.J) GO TO 3
      DIST=SQRT((X(I)-X(J))**2+(Y(I)-Y(J))**2)
      IF(DIST.LT.DMIN) DMIN=DIST
    3 CONTINUE
      DMEAN=DMEAN+DMIN
    2 CONTINUE
      DMEAN=DMEAN/FN
      WRITE(6,140) DMEAN
  140 FORMAT(//' OBSERVED MEAN DISTANCE = ',F10.4)
C
C     CALCULATE THE EXPECTED MEAN DISTANCE
C
      EXPECT=1/(2*SQRT(FN/A))
      WRITE(6,150) EXPECT
  150 FORMAT(/' EXPECTED MEAN DISTANCE = ',F10.4)
C
C     CALCULATE THE NEAREST NEIBOR STATIATIC
C
      R=DMEAN/EXPECT
      WRITE(6,160) R
  160 FORMAT(/' NEAREST NEIBOR STATISTIC = ',F8.4)
C
C     SIGNIFICANCE TEST
C
      SE=(0.26136*SQRT(A))/FN
```

付図1（続き）

```
      Z=(DMEAN-EXPECT)/SE
      WRITE(6,170) SE
  170 FORMAT(/' STANDARD ERROR = ',F10.4)
      WRITE(6,180) Z
  180 FORMAT(/' Z-SCORE = ',F8.4)
      IF(ABS(Z).GE.2.576) GO TO 4
      IF(ABS(Z).GE.1.96) GO TO 5
      WRITE(6,190)
  190 FORMAT(/' NEAREST NEIBOR STATISTIC IS NOT SIGNIFICANT ')
      GO TO 6
    4 WRITE(6,200)
  200 FORMAT(/' NEAREST NEIBOR STATISTIC IS SIGNIFICANT AT THE 0.01 LEVE
     1L FOR A TWO-TAILED TEST ')
      GO TO 6
    5 WRITE(6,210)
  210 FORMAT(/' NEAREST NEIBOR STATISTIC IS SIGNIFICANT AT THE 0.05 LEVE
     1L FOR A TWO-TAILED TEST ')
    6 STOP
      END
```

```
NUMBER OF POINTS =     19       AREA =     1840.0 KM**2

        X-COOR    Y-COOR

   1     17.90     32.50
   2     34.30     34.00
   3     13.10     29.10
   4     23.10     30.30
   5     40.20     29.60
   6      7.90     24.90
   7     17.60     24.00
   8     30.60     25.40
   9     26.30     21.50
  10     40.50     20.70
  11      7.70     18.30
  12     18.80     13.30
  13     32.10     18.80
  14     14.30      9.60
  15     27.90     12.60
  16     39.40     16.60
  17      6.20      7.50
  18     22.00      7.20
  19     35.20     10.40

OBSERVED MEAN DISTANCE =     6.3308

EXPECTED MEAN DISTANCE =     4.9204

NEAREST NEIBOR STATISTIC =   1.2866

STANDARD ERROR =      0.5901

Z-SCORE =   2.3902

NEAREST NEIBOR STATISTIC IS SIGNIFICANT AT THE 0.05 LEVEL FOR A TWO-TAILED TEST
```

付図 2　空間的連関係数の FORTRAN プログラムと出力例

```
C
C ************************************************************
C *      PROGRAM OF COEFFICIENT OF SPATIAL ASSOCIATION        *
C *                                                           *
C *       (INPUT DATA)                                        *
C *          N1: NUMBER OF POINTS IN DISTRIBUTION TYPE 1      *
C *          N2: NUMBER OF POINTS IN DISTRIBUTION TYPE 2      *
C *            -NOTE- IF TWO DISTRIBUTIONS HAVE THE SAME      *
C *                   NUMBER OF POINTS, LEAVE N2 BLANK.       *
C *                                                           *
C *          X1,Y1: X-,Y-COORDINATES OF DISTRIBUTION TYPE 1   *
C *          X2,Y2: X-,Y-COORDINATES OF DISTRIBUTION TYPE 2   *
C ************************************************************
C
      DIMENSION X1(200),Y1(200),X2(200),Y2(200)
C
C     READ INPUT DATA
C
      READ(5,100) N1,N2
  100 FORMAT(2I5)
      IF(N2.NE.0) GO TO 1
      N2=N1
    1 WRITE(6,110)
  110 FORMAT(' INPUT DATA OF DISTRIBUTION 1 '//8X,'X1-COOR   Y1-COOR'/)
      DO 2 I=1,N1
      READ(5,120) X1(I),Y1(I)
      WRITE(6,130) I,X1(I),Y1(I)
  120 FORMAT(2F10.2)
  130 FORMAT(I5,2F10.2)
    2 CONTINUE
      WRITE(6,140)
  140 FORMAT(/' INPUT DATA OF DISTRIBUTION 2 '//8X,'X2-COOR   Y2-COOR'/)
      DO 3 I=1,N2
      READ(5,120) X2(I),Y2(I)
      WRITE(6,130) I,X2(I),Y2(I)
    3 CONTINUE
C
C     CALCULATE THE NEAREST NEIGHBOR DISTANCES
C
      CALL MD1(N1,X1,Y1,DMN1)
      CALL MD1(N2,X2,Y2,DMN2)
      CALL MD2(N1,N2,X1,Y1,X2,Y2,DMN3)
      CALL MD2(N2,N1,X2,Y2,X1,Y1,DMN4)
C
C     CALCULATE COEFFICIENT OF SPATIAL ASSOCIATION
C
      SA=((DMN1+DMN2)-(DMN3+DMN4))/((DMN1+DMN2)+(DMN3+DMN4))
      WRITE(6,150) SA
  150 FORMAT(///' COEFFICIENT OF SPATIAL ASSOCIATION = ',F8.5)
      STOP
      END
```

付図2（続き・その1）

```
C
C     SUBROUTINE FOR SUM OF INTRA-DISTRIBUTION N-N DISTANCES
C
      SUBROUTINE MD1(N,X,Y,DMN)
      DIMENSION X(200),Y(200),DM(200)
      DO 1 I=1,N
      DM(I)=9000000.0
      DO 2 J=1,N
      IF(I.EQ.J) GO TO 2
      DX=X(I)-X(J)
      DY=Y(I)-Y(J)
      DIST=SQRT(DX*DX+DY*DY)
      IF(DIST.LT.DM(I)) DM(I)=DIST
    2 CONTINUE
    1 CONTINUE
      DMN=0.
      DO 3 I=1,N
      DMN=DMN+DM(I)
    3 CONTINUE
      RETURN
      END
```

```
C
C     SUBROUTINE FOR SUM OF INTER-DISTRIBUTION N-N DISTANCES
C
      SUBROUTINE MD2(N,M,X1,Y1,X2,Y2,DMN)
      DIMENSION X1(200),X2(200),Y1(200),Y2(200),DM(200)
      DO 1 I=1,N
      DM(I)=9000000.0
      DO 2 J=1,M
      DX=X1(I)-X2(J)
      DY=Y1(I)-Y2(J)
      D=DX*DX+DY*DY
      IF(D.EQ.0.) GO TO 3
      DIST=SQRT(D)
      IF(DIST.LT.DM(I)) DM(I)=DIST
    2 CONTINUE
      GO TO 1
    3 DM(I)=0.
    1 CONTINUE
      DMN=0.
      DO 4 I=1,N
      DMN=DMN+DM(I)
    4 CONTINUE
      RETURN
      END
```

付図2（続き・その2）

INPUT DATA OF DISTRIBUTION 1

	X1-COOR	Y1-COOR
1	17.90	32.50
2	34.30	34.00
3	13.10	29.10
4	23.10	30.30
5	40.20	29.60
6	7.90	24.90
7	17.60	24.00
8	30.60	25.40
9	26.30	21.50
10	40.50	20.70
11	7.70	18.30
12	18.80	13.30
13	32.10	18.80
14	14.30	9.60
15	27.90	12.60
16	39.40	16.60
17	6.20	7.50
18	22.00	7.20
19	35.20	10.40

INPUT DATA OF DISTRIBUTION 2

	X2-COOR	Y2-COOR
1	17.90	32.70
2	32.10	34.10
3	11.30	27.90
4	25.50	29.20
5	39.70	30.70
6	4.50	23.30
7	18.80	24.70
8	32.80	25.90
9	26.30	21.20
10	40.50	22.60
11	5.50	15.30
12	19.60	16.40
13	33.70	17.90
14	12.90	12.00
15	27.10	13.30
16	41.30	14.60
17	6.20	7.20
18	20.40	8.60
19	34.30	9.90

COEFFICIENT OF SPATIAL ASSOCIATION = 0.57481

付図3　ユークリッド2次元回帰分析の FORTRAN プログラムと出力例

```
      C
      C ***********************************************************
      C *   PROGRAM OF EUCLIDEAN BIDIMENSIONAL REGRESSION ANALYSIS   *
      C *                                                           *
      C *     (INPUT DATA)                                          *
      C *         N   : NUMBER OF DATA POINTS                       *
      C *         X,Y: INDEPENDENT COORDINATES                      *
      C *         U,V: DEPENDENT COORDINATES                        *
      C ***********************************************************
      C
            DIMENSION X(300),Y(300),U(300),V(300),UP(300),VP(300),UE(300),
           1VE(300),S(4),S2(4),C(4,4),VAR(4),B(4)
      C
      C     READ INPUT DATA
      C
            READ(5,100) N
        100 FORMAT(I5)
            DO 1 I=1,N
            READ(5,110) X(I),Y(I),U(I),V(I)
        110 FORMAT(4F10.3)
          1 CONTINUE
      C
            DO 2 I=1,4
            S(I)=0.0
            S2(I)=0.0
            DO 2 J=1,4
            C(I,J)=0.0
          2 CONTINUE
            FN=FLOAT(N)
      C
      C     CALCULATE MEAN
      C
            DO 3 I=1,N
            S(1)=S(1)+X(I)
            S(2)=S(2)+Y(I)
            S(3)=S(3)+U(I)
            S(4)=S(4)+V(I)
          3 CONTINUE
            DO 4 I=1,4
            B(I)=S(I)/FN
            S(I)=0.0
          4 CONTINUE
      C
      C     CALCULATE SUM OF SQUARES AND SUM OF PRODUCTS
      C
            DO 5 I=1,N
            UP(I)=X(I)-B(1)
            VP(I)=Y(I)-B(2)
            UE(I)=U(I)-B(3)
            VE(I)=V(I)-B(4)
            S(1)=S(1)+UP(I)
            S(2)=S(2)+VP(I)
            S(3)=S(3)+UE(I)
            S(4)=S(4)+VE(I)
            S2(1)=S2(1)+UP(I)*UP(I)
            S2(2)=S2(2)+VP(I)*VP(I)
            S2(3)=S2(3)+UE(I)*UE(I)
            S2(4)=S2(4)+VE(I)*VE(I)
```

付図３（続き・その１）

```
      C(1,2)=C(1,2)+UP(I)*VP(I)
      C(1,3)=C(1,3)+UP(I)*UE(I)
      C(1,4)=C(1,4)+UP(I)*VE(I)
      C(2,3)=C(2,3)+VP(I)*UE(I)
      C(2,4)=C(2,4)+VP(I)*VE(I)
      C(3,4)=C(3,4)+UE(I)*VE(I)
    5 CONTINUE
C
C     ESTIMATE TRANSFORMATION EQUATIONS
C
      DO 6 I=1,4
      VAR(I)=S2(I)/FN
    6 CONTINUE
      SUM=S2(1)+S2(2)
      A11=(C(1,3)+C(2,4))/SUM
      A12=(C(2,3)-C(1,4))/SUM
      A21=-A12
      A22=A11
      A13=B(3)-A11*B(1)-A12*B(2)
      A23=B(4)-A21*B(1)-A22*B(2)
      S2(1)=0.0
      S2(2)=0.0
      WRITE(6,120)
  120 FORMAT(' EUCLIDEAN TRANSFORMATION'/)
      WRITE(6,130)
  130 FORMAT(8X,'X-COOR',4X,'Y-COOR',4X,'U-COOR',4X,'V-COOR',3X,'UP-COOR
     1'3X,'VP-COOR'/)
      DO 7 I=1,N
      UP(I)=A13+A11*X(I)+A12*Y(I)
      VP(I)=A23+A21*X(I)+A22*Y(I)
      UE(I)=U(I)-UP(I)
      VE(I)=V(I)-VP(I)
      D1=UE(I)*UE(I)
      D2=VE(I)*VE(I)
      S2(1)=S2(1)+D1
      S2(2)=S2(2)+D2
C
C     WRITE COORDINATE INFORMATION
C
      WRITE(6,140) I,X(I),Y(I),U(I),V(I),UP(I),VP(I)
  140 FORMAT(I4,6F10.2)
    7 CONTINUE
C
C     WRITE GOODNESS-OF-FIT AND TRANSFORMATION INFORMATION
C
      DETZ=SQRT(1.0-(S2(1)+S2(2))/(S2(3)+S2(4)))
      FPCT=DETZ*DETZ*100.
      DET=A11*A22-A12*A21
      IF(DET.EQ.0.) WRITE(6,150)
  150 FORMAT(' ZERO DETERMINANT, WATCH OUT')
      IF(DET.EQ.0.) GO TO 8
      WRITE(6,160) DETZ,FPCT,DET
  160 FORMAT(//' (GOODNESS-OF-FIT) '/' R = ',F9.4,'  PERCENT FIT = ',
     1F8.2/' DETERMINANT = ',F10.2)
      WRITE(6,170)
  170 FORMAT(//' (TRANSFORMATION EQUATIONS)')
      WRITE(6,180) A13,A11,A12,A23,A21,A22
  180 FORMAT(' A1 = ',E15.6,'      B1 = ',E15.6,'     -B2 = ',E15.6/
     1        ' A2 = ',E15.6,'      B2 = ',E15.6,'      B1 = ',E15.6)
    8 STOP
      END
```

付図3（続き・その2）

EUCLIDEAN TRANSFORMATION

	X-COOR	Y-COOR	U-COOR	V-COOR	UP-COOR	VP-COOR
1	17.90	32.70	17.90	32.50	18.53	32.19
2	32.10	34.10	34.30	34.00	32.16	33.37
3	11.30	27.90	13.10	29.10	12.15	27.67
4	25.50	29.20	23.10	30.30	25.77	28.75
5	39.70	30.70	40.20	29.60	39.40	30.03
6	4.50	23.30	7.90	24.90	5.57	23.34
7	18.80	24.70	17.60	24.00	19.30	24.51
8	32.80	25.90	30.60	25.40	32.73	25.51
9	26.30	21.20	26.30	21.50	26.45	21.07
10	40.50	22.60	40.50	20.70	40.08	22.25
11	5.50	15.30	7.70	18.30	6.44	15.66
12	19.60	16.40	18.80	13.30	19.97	16.55
13	33.70	17.90	32.10	18.80	33.51	17.83
14	12.90	12.00	14.30	9.60	13.50	12.41
15	27.10	13.30	27.90	12.60	27.13	13.49
16	41.30	14.60	39.40	16.60	40.75	14.58
17	6.20	7.20	6.20	7.50	7.02	7.88
18	20.40	8.60	22.00	7.20	20.65	9.06
19	34.30	9.90	35.20	10.40	33.99	10.15

(GOODNESS-OF-FIT)
R = 0.9884 PERCENT FIT = 97.69
DETERMINANT = 0.92

(TRANSFORMATION EQUATIONS)
A1 = 0.994443E+00 B1 = 0.958656E+00 -B2 = 0.114248E-01
A2 = 0.105087E+01 B2 = -0.114248E-01 B1 = 0.958656E+00

POINT PATTERN ANALYSIS AND ITS FORTRAN PROGRAMS

Yoshio Sugiura*

*Center for Urban Studies, Tokyo Metropolitan University

Comprehensive Urban Studies, Special Edition, 1989, pp. 5 −23

This paper aims at introducing some methods of point pattern analysis. Algorithms of nearest neighbor measure (Clark and Evans, 1954), coefficient of spatial association (Sorensen, 1974) and Euclidean bidimensional regression analysis (Tobler, 1965, 1983) are described briefly. Their FORTRAN programs are presented in the appendix.