GEOGRAPHIA POLONICA 27, 1973

AN ANALYSIS OF THE TYPES OF SPATIAL DISTRIBUTION OF TOWNS

Zbyszko Chojnicki and Teresa Czyż

An analysis of the distribution of geographical objects in space is an important step in the generalization procedures of economic geography, since the purpose of this analysis is to explain the occurrence and variation of spatial patterns. Such generalizations are codetermined by multi-variate conditions; hence they are as a rule of an approximate character, and therefore the researcher's principal effort must focus on increasing the accuracy of the approximations. The formulation of generalizations is based upon hypotheses, i.e. tentative solutions that have to be tested. A hypothesis with a high level of confirmation may be included in the theory. The formulation and systematic testing of hypotheses is thus an indispensable step in the generalization procedure.

Recent years have witnessed important advances in this approach, mainly due to the construction of mathematical models of the different spatial patterns which contain definite assumptions about the processes that govern the patterns rather than mere statistical description. The principal methodological problem here is to find models that identify the processes governing the distribution of the objects studied. Stochastic models have proved particularly useful in explaining the distribution and changes in the spatial patterns of human activities.

Such models can be developed by making certain assumptions concerning the mathematical processes that generate some types of probability distributions. These mathematical processes may in some cases refer directly to the processes that occur in reality. Thus a number of probability distributions can be employed as models of geographical processes. The Poisson distribution is among others particularly useful for the analysis of the spatial distribution of points. It occurs in various forms depending on the concrete phenomena to which it may be applied. If a set of events or objects is randomly distributed in space (or time), the probability of an event or object occurring in any part of that area (or in any time interval) is defined by the Poisson distribution.

J. Coleman (1964, p. 291) states, that "... the Poisson process is appropriate to social phenomena because it constitutes a rational model whose assumptions can mirror our assumptions about actual phenomena". The usefulness of the Poisson process in the analysis of social processes is mainly due to the assumptions on which it is based. Firstly, it describes a certain number of elements (or proportions) and events. In this case, therefore, it is not necessary to perform measurements of the values of continuous variables, which occur very rarely in the social sciences anyway. Secondly, the phenomena described by the Poisson distribution are much more continuous in time (and space) than, say, these described by the binomial distribution, which is more of a description of series of discrete phenomena. Hence the considerable usefulness of the Poisson distribution in studies of spatial distribution; it opens up perspectives for the development of "genetic probability" as the fundamental language for discussing different geographical forms.

The Poisson distribution and other related distributions have been employed as models in a number of mathematical representations in analyses of spatial patterns. These compare the observed pattern with the theoretical one generated by such distributions, and allow measurement of deviations from the particular random processes. These methods include for example D. Harvey's (1966) quadratic sampling, the measures of contiguity of M. Dacey (1965) and A. Cliff (1968), the method of the nearest neighbour of M. Dacey (1962), and A. Getis' (1967) sequence analysis.

The employment of these methods in geographical analyses shows how useful they may be in constructing and testing geographical hypotheses. In studies in settlement geography, the construction and testing of hypotheses through stochastic models involves mainly the description and analysis of the spatial distributions representing patterns of settlements location.[1] The present study is devoted to such an analysis of the spatial distribution of towns and cities using the nearest neighbour method. This method, devised by I. Matui (1932) in his classical study utilizing the Poisson distribution, was later developed by P. J. Clark and F. C. Evans (1954) in the field of plant ecology, and by M. Dacey (1960, 1962) and by J. V. Medvedkov (1967) in settlement geography.

Hypotheses concerning the form of the distribution of the particular points of the settlements network are based on empirical data or theoretical premises. In the former case hypotheses are formulated on the strength of observations, i.e., of a description made in terms of numerical characteristics or maps; this is particularly difficult when many variables are involved. In the latter case, theoretical premises may be derived from different theories. The classical model of constructing and testing hypotheses relating to the spatial distribution of towns is usually taken as Christaller's theory of central places (1933).

Christaller was the first to attempt the construction of a theoretical model of the distribution of towns. He tried to explain the location pattern of towns in terms of the functions performed by the individual towns on behalf of their surrounding areas. The theory of central places, which Christaller derived from an analysis of the market reach of commodities and services, by assuming that all constituents of the populated area can be supplied in terms of the smallest possible number of central settlements, is subject to geometrical laws and constitutes a hexagonal lattice. The deviations from the ideal pattern, based on the principle of supply, account for two further principles — those of communication and administration. Christaller's theory was severely criticized for its statical approach to the problem, and for the narrowly model-oriented assumptions which saw the spatial order of economic phenomena dependent only on some service functions.

The criticism of Christaller's theory need not detain us here, but it may be remarked that the picture of the spatial order furnished by this theory is only indirectly reflected in reality. This is presumably due to the predominantly hypothetical and deductive character of the theory. Hence its laws can explain reality only to the extent to which its model premises are realized in it. Apart from this however, it is to the theory of central places that we owe the development of a rational model of the spatial distribution of towns which may also follow from other premises. The recognition of that order represented by the hexagonal lattice is not in itself a direct test of Christaller's theory, yet it casts some light upon the nature of the processes governing the actual order in geographical space, represented by a pattern which may be regular (hexagonal), random or clustered.

[1] The operational definition of the pattern of distribution (location) of settlements treats them as a set of material points over a definite area.

To identify a distributional pattern of towns as a definite type, it would seem promising to start by identifying randomness in the distribution of points on a plane, using the nearest neighbour distance technique.[2]

An alternative approach to the description of point patterns is Dacey's regional method of analysis of the nearest neighbour. The mathematical description of this method is contained in the papers of M. Dacey (1960, 1962), M. Dacey and T. Tung (1962). This method is used with increasing frequency in geographical studies of point patterns.

Our empirical study also illustrates, the use of the nearest neighbour method in describing properties of distribution of Polish towns both in different spatial terms and for different categories of towns, namely for three types of distributions:

(1) for *poviat* towns on the national scale,
(2) for *poviat* towns on the voivodship scale,
(3) for all towns of the Poznań voivodship.

The change of spatial scale and of the categories of towns was intended to detail the specific character of the distribution.

TABLE 1. Nearest neighbour statistics for the set of *poviat* towns of Poland

Sector k	Observed mean	Mean of distribution which is			Ratio of randomness
		hexagonal	random	clustered	
	$\bar{d}_{ik}$	E_h	$E_{rk/6}$	E_c	R_k
1	20.5	38.05	17.85	1.0	1.148
2	26.3	38.05	28.08	1.0	0.937
3	32.6	38.05	37.15	1.0	0.851
4	37.1	38.05	46.55	1.0	0.797
5	43.0	38.05	57.76	1.0	0.744
6	53.1	38.05	75.03	1.0	0.708
		D	D	D	R
		27.2	26.8	88.0	0.806

The analysis of the *poviat* towns for the whole of Poland comprised 264 towns with a density of 0.0008 per 1 km². Their mean nearest neighbour distance is 20.5 km, and the mean 6-sectoral distance, 35.2 km. The results of measurements and calculations, that is the actual mean values and the mean values of the hexagonal and the random and the clustered distributions, make it possible both to analyse the spatial distribution of points and to test the hypothesis that the towns constitute a hexagonal pattern rather than either of the two alternative random or the clustered patterns.[3] The mean values indicate that

[2] J. O. Abiodun (1967) used factor analysis to test Christaller's model of central places. She showed that the distribution of the centres of the particular orders in the regional system of the Ijebu Province (Nigeria) is on the whole in accordance with Christaller's model for $k = 3$.

[3] The random distribution of points on a plane is identified by the set of mean theoretical regional distances to the neighbours generated from the Poisson distribution.

the observed pattern is neither fully hexagonal, random or clustered, because the sets of mean observed and mean expected values do not exactly coincide with each other (Table 1).

To provide a basic test for stating which of the three theoretical patterns best fits the actual one, let us assume that the best agreement between the observed and the theoretical patterns is when the differences between the corresponding mean values are smallest; this can be expressed by the formula:

$$D^2 = \sum_{k=1}^{k=K} [\bar{d}_{tk} - E_k]^2$$

where E_k is the expected (theoretical) value for sector k. In the above analysis the value D is smallest for the random pattern (26.8); this enables us to reject the hypothesis that the *poviat* towns in Poland constitute a hexagonal pattern. The total ratio of randomness amounts to 0.806. The ratios of randomness for the individual sectors decrease with the growth of k (1.148–0.708), which suggests that the first neighbours are more uniformly distributed than the more distant ones (the value of $R = 1.148$ suggesting a trend towards a hexagonal distribution).

TABLE 2. Nearest neighbour statistics for the *poviat* towns of the voivodships

Voivodship	Number of towns	Density of towns	Observed mean sectoral distance	Expected mean sectoral distance in the random distribution	Ratio of randomness
Białystok	8	0.00094	42.1	40.8	1.032
Bydgoszcz	21	0.00106	35.5	39.5	0.899
Gdańsk	8	0.00106	34.2	38.3	0.893
Katowice	22	0.00275	21.1	23.5	0.895
Kielce	21	0.00108	35.6	38.3	0.930
Koszalin	9	0.00072	42.8	45.4	0.943
Cracow	12	0.00139	30.3	33.1	0.915
Lublin	14	0.00088	39.8	42.2	0.944
Łódź	20	0.00115	35.1	37.1	0.946
Olsztyn	14	0.00844	41.4	42.2	0.982
Opole	10	0.00142	31.9	33.2	0.966
Poznań	30	0.00111	36.2	37.1	0.971
Rzeszów	13	0.00168	32.1	29.9	1.071
Szczecin	5	0.00098	37.4	39.5	0.948
Warsaw	29	0.00101	37.1	39.5	0.937
Wrocław	16	0.00129	31.1	34.9	0.890
Zielona Góra	12	0.00113	35.6	37.1	0.958

The distribution of *poviat* towns in the voivodships show ratios of randomness between 1.071 and 0.890, plus the lowest values of the criterion D for the random distribution, which suggests that the distribution of towns on this scale is also random (Table 2). The ratios of randomness for the voivodships show some differentiation, but voivodships tending toward the hexagonal pattern (those of Rzeszów and Białystok) can nevertheless be distinguished from those

tending toward the clustered pattern (Katowice, Gdańsk and Wrocław) (Fig. 1). It must be stressed though, that the ratios of randomness for the first sector are as a rule relatively high (the maximum being reached in the Szczecin voivodship — 1.469) and that the ratio drops below 1 (0.947) only in the case of the Katowice voivodship.

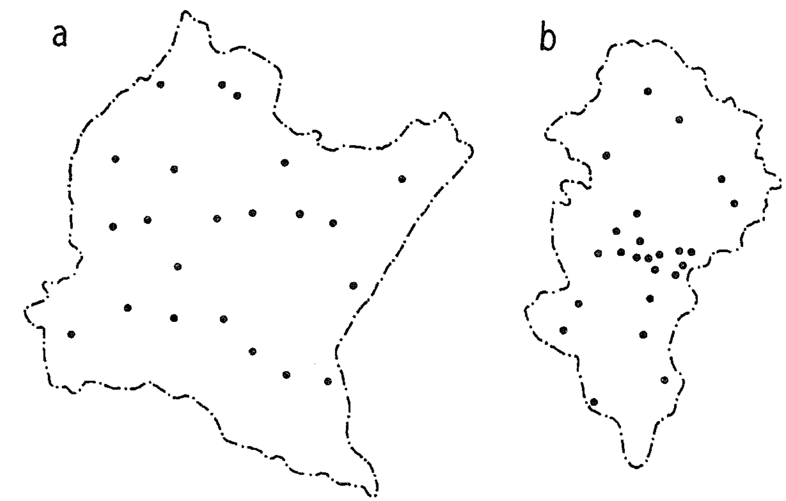

Fig. 1. a — Rzeszów voivodship, 1 b — Katowice voivodship

If the distribution of towns is considered to be a result of the effects of different socio-economic, physico-geographic and other factors, the ideal theoretical distribution is treated as a background and setting for the empirical distribution.

TABLE 3. Correlation between the ratio of randomness and the socio-economic variables

Variable	Pearson correlation coefficient
Population density per 1 km²	— 0.4319
Share of urban in total population (%)	— 0.4658
Value of agricultural output per 100 ha of agricultural land	— 0.5372
Industrial employment per 100 km²	— 0.8818
Public roads per 100 km²	— 0.5479

bution. Interpretations of deviations from the uniform distribution which can be seen in the various deviations and distortions is based upon an analysis of the correlations between the value of the ratio R and the socio-economic variables (Table 3).

Ratio of randomness shows a very good negative correlation with the feature "industrial employment per 100 km²" (—0.88), which is regarded as statistically significant at the $\alpha = 0.01$ level. The value of the coefficient of determination suggests that this variable explains 77% of the variation of the ratio of randomness (as calculated for 17 voivodships). Thus the observed distribution of towns is conditioned by industrial employment, an important factor in the formation of the clustered pattern (the voivodship of Katowice being one example) The remaining variables of population density, urban population, agricultural output and road network density, are also negatively correlated with the ratio of randomness, though not significantly. Since the analysed factors constitute a group of variables with the closest possible links, the degree of industrialization can be regarded as the most important factor in the observed deviations from the hexagonal pattern.

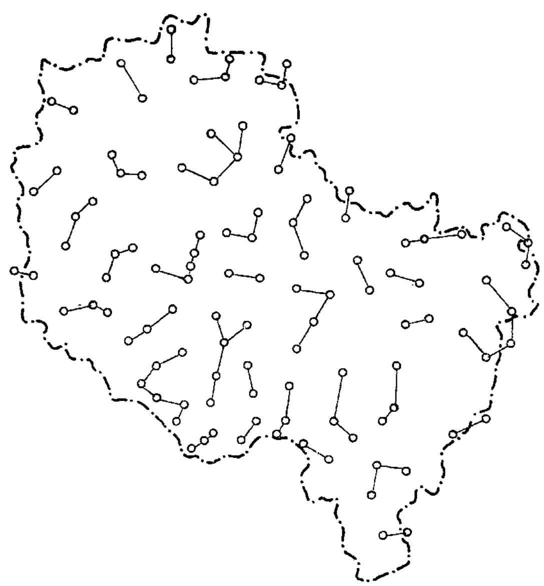

Fig. 2. Poznań voivodship

The observed distances of the towns of the Poznań voivodship are between 4.0 and 50.1 km. The towns exhibit no tendency to cluster, the most frequent form being isolated 2–3-element subsets (Fig. 2). The mean regional nearest neighbour distances are between 11.1 and 30.0 km, increasing systematically in each sector by 3 km except the last, when the change was 6.6 km.

The observed distances to the first nearest neighbour for the set of 102 towns show no correlation (correlation coefficient 0.07) with the population number. Thus the distribution of towns in the Poznań voivodship cannot be explained by this factor (as done by E. N. Thomas, 1961 and J. V. Medvedkov, 1963).

The analysis of the relation between the type of town, defined on the basis of the population's occupational structure, and its nearest neighbour distance also produced unsatisfactory results. No regularities were found in the increase or decrease of the distance to the nearest neighbour depending on the function of the town. However, the relatively small mean distances to the nearest neighbour for the agricultural towns (10.5 km), and the relatively large ones for service and industrial-service towns (13.2 and 14.5 km) were found. Presumably, if we used another method using the proportions of the exogenous group to define the specialization of functions we would get different results, confirm-

ing the hypothesis that the trend toward specialization of functions is inversely proportional to the distance to other urban centres.

The ratio of randomness is 0.971, which with the smallest value D for the random distribution, indicates a random pattern in towns of the Poznań voivodship (Table 4).

TABLE 4. Nearest neighbour statistics for the towns of the Poznań voivodship

Sector	Observed mean	Mean of distribution which is			Ratio of randomness
		hexagonal	random	clustered	
	$\bar{d}_{ik}$	E_h	$E_{rk/6}$	E_c	R_k
1	11.1	17.4	8.2	1.0	1.354
2	14.4	17.4	12.9	1.0	1.116
3	17.7	17.4	17.0	1.0	1.041
4	20.4	17.4	21.4	1.0	0.953
5	23.4	17.4	26.5	1.0	0.883
6	30.0	17.4	34.4	1.0	0.872
		D	D	D	R
		15.8	6.4	47.7	0.971

This ratio shows little local variation, which can be seen from the more or less equal values for all the general-economic regions of the Poznań voivodship (Table 5).

TABLE 5. Nearest neighbour statistics for the towns of the particular general-economic regions of the Poznań voivodship

Region	Number of towns	Density of towns	Observed mean sectoral distance	Expected mean sectoral distance in random distribution	Ratio of randomness
Northern	9	0.003169	21.1	21.9	0.9648
Central	46	0.003693	19.2	20.1	0.9587
Eastern	15	0.003531	21.3	20.7	1.0254
South-eastern	20	0.003699	19.9	20.4	0.9766
South-western	12	0.006030	15.9	15.9	1.0000

The present study attempts to test hypotheses concerning the spatial distribution of Polish towns by using statistics of distances between the towns in the nearest neighbour method. Our analysis has shown the distribution of towns to be a random one irrespective of the scale applied. The characteristics of these distributions was based upon a hypothetical mathematical process of the Poisson distribution type. Attempts at interpreting this random distribution in geographical terms prove rather difficult because we still do not know exactly which factors are responsible for a given type of distribution. That the distribution of Polish towns forms neither a hexagonal pattern nor exhibits any distinct clustering may perhaps be due to the fact that the broad geographic processes which determine the actual spatial distribution of the towns, do not merely result from the premises assumed in the theory of central places, but are governed by much more complex multi-factor genetic-functional mechanisms. One significant factor tending the random distribution to develop into the clustered pattern is the process of industrialization.

Further progress in the assessment and interpretation of the deformations of the pattern must be based both on the study of different sizes of towns and functional types, and on modifying the Poisson distribution as the mathematical model.

Adam Mickiewicz University, Poznań

BIBLIOGRAPHY

Abiodun, J. O., 1967, Urban hierarchy in a developing country, *Econ. Geogr.*, 45, 347–367.

Christaller, W., 1933, *Die zentralen Orte in Süddeutschland*, Jena.

Clark, P. J., Evans, F. C., 1954, Distance to nearest neighbour, as a measure of spatial relationships in population, *Ecology*, 35, 445–453.

Cliff, A., 1968, The neighbourhood effect in the diffusion of innovations, *Transactions, Inst. Brit. Geogr.*, 44, 75–84.

Coleman, J., 1964, *Introduction to mathematical sociology*, New York.

Dacey, M., 1960, A note on the derivation of nearest-neighbor distances, *J. Reg. Sci.*, 2, 81–87.

Dacey, M., 1962, Analysis of central place and point pattern by a nearest neighbour method, in: K. Norborg (ed.), *Proceedings of the IGU Symposium in Urban Geography*, Lund 1960, Lund Studies in Geography, Ser. B, Human Geogr., 24, 55–75.

Dacey, M., Tung, T., 1962, The identification of randomness in point patterns, *J. Reg. Sci.*, 4, 83–96.

Dacey, M., 1965, *Measures of contiguity for k coloured maps*, Dep. of Geogr., Northwestern Univ.

Getis, A., 1967. A method for the study of sequences in geography, *Transactions, Inst. Brit. Geogr.*, 42, 87–92.

Harvey, D., 1966, Geographical processes and the analysis of spatial point patterns, *Transactions, Inst. Brit. Geogr.*, 40, 81–95.

Matui, I., 1932, Statistical study of the distribution of scattered villages in two regions of the Tonami Plain, Toyama Prefecture, *Jap. J. Geol. Geogr.*, 9, 251–256.

Medvedkov, J. V., 1967, The concept of entropy in settlement pattern analysis, *Reg. Sci. Ass., Papers*, 18, 165–168.

Thomas, E. N., 1961, Toward an explanded central place model, *Geogr. Rev.*, 51, 400–411.