QUADRAT ANALYSIS OF URBAN DISPERSION

by

Andrei Rogers

Working Paper No. 93

This research has been partially supported by a grant from the Economic Development Administration.

March, 1969

PREFACE

This paper is the third working paper of a group that have as their common concern the development of statistical methods for the analysis of the spatial structure of population and employment in urban areas and regions. Its principal concern is the application of quadrat methods to the analysis of urban dispersion.

While conducting this study over the past year, I have been assisted by many people. In particular, I wish to acknowledge my debt to three able research assistants: Norbert Gomar, Juan Martin and Jorge Meisse. My thanks also go to Lidija Podbregar-Vasle of the Town Planning Institute of Slovenia, Yugoslavia and to Noreen Roberts of the Bay Area Transportation Study staff for providing me with the data on the spatial patterns of retail establishments in Ljubljana and San Francisco, respectively.

This work is being conducted under the sponsorship of a grant from the Economic Development Administration of the U.S. Department of Commerce. A complete listing of the Working Papers under this grant appears on the back page of this paper. They deal with regional patterns, development, and policies.

Andrei Rogers
Center for Planning and
Development Research
University of California, Berkeley

Working Papers on Spatial Pattern Analysis

#86 Andrei Rogers and Norbert G. Gomar, "Statistical Inference in Quadrat Analysis," October, 1968, to be published in Geographical Analysis.

#91 Shlomo Angel, "Point Patterns in Urban Regions: Two Complementary Tests," January, 1969.

#93 Andrei Rogers, "Quadrat Analysis of Urban Dispersion," March, 1969.

Andrei Rogers and Juan Martin, "Bivariate Quadrat Analysis of the Dispersion of Employment and Population in Urban Areas," forthcoming.

QUADRAT ANALYSIS OF URBAN DISPERSION

PART I

INTRODUCTION

1.1 Introduction

In metropolitan regions across the world, efforts to cope with the pressing problems of growth and development have pointed to the need for a fuller understanding of the spatial structure of urban areas. Scholars in various disciplines have sought theories and methods which might contribute to an improved comprehension of the organization and distribution of human activities in space, their interlinkages, and the flows of people, goods and information between them. Such endeavors have been supported with the belief that a better understanding of city-systems will lead to a firmer foundation on which to base policy decisions that are designed to effect desired changes in them.

The geographical arrangement of activities in urban areas has particularly received a great deal of attention from economists, geographers, city planners and ecologists. Despite their efforts, however, we still do not have well-developed and tested methods which would allow us to quantitatively express the spatial patterns of various classes of urban activities and to compare them on an intraurban and an interurban level. Recent efforts by quantitative geographers and plant ecologists, however, suggest that, ultimately, a body of techniques which provide such quantifications and comparisons may evolve out of a methodology that commonly is referred to as spatial point pattern analysis. A subclass of that methodology, called quadrat analysis, is the concern of this paper.

1.2 The Statistical Analysis of Dispersion

1.2.1 Definitions

In recent years, location theorists have become increasingly aware of the need for quantitative descriptions of two-dimensional point patterns. Geographers, in particular, have recognized that pattern is the geometrical expression of location theory and that techniques of pattern analysis, therefore, form an important methodological component of urban geography [Hudson and Fowler (1966)].

The development of methods for describing and analyzing spatial patterns has been hampered by the lack of a precise definition of what pattern is and how its properties differ from those of shape and dispersion. Every spatial arrangement of objects possesses these three properties and may be uniquely defined in terms of them. Shape is a two-dimensional characteristic of a spatial arrangement that is defined by a closed curve [Bunge (1962)] which delineates the collection of objects and provides an areal measure of their distribution. Pattern is a zero-dimensional characteristic of a spatial arrangement which describes the relative spacing of a set of objects with respect to one another [Hudson and Fowler (1966)]. Finally, dispersion may be viewed as a one-dimensional characteristic of a spatial arrangement which measures the relative spacing of a set of objects with respect to one particular shape of a given area [McConnell (1966)]. Thus we may view dispersion as an attribute of a pattern that is located within a particular shape, at a given density.

For a graphical illustration of the above distinctions between shape, pattern and dispersion, let us turn to Figure 1.1 below. Figure 1.1a provides an example of two identical shapes having different areas; Figure 1.1b

illustrates two identical patterns having different modules;[1] and Figure 1.1c demonstrates that two identical patterns placed into identical shapes, can lead to identical dispersions. Figures 1.1d, 1.1e, and 1.1f, respectively, illustrate converse examples, i.e., different shapes with identical areas, different patterns with identical modules, and different dispersions with identical shapes and densities.

1.2.2 The Random Spatial Point Pattern

A fundamental concept which is used throughout this paper is the notion of "randomness." Since we shall frequently refer to this concept, it is necessary to define it unequivocally and in mathematical terms.

A random spatial point pattern is defined to be the geographical disposition of points on a plane that is generated by a mechanism, or, more formally, a realization of a spatial point process, that satisfies the following two conditions:

1. Condition of Equal Probability--Any point has an equal probability of occurring at any position in the plane, and any subregion of the plane, therefore, has an equal probability of containing a point as any other subregion of equal area.

2. Condition of Independence--The position of a point in the plane is independent of the position of any other point.

Any point located in accordance with the above two conditions is said to be located randomly and the pattern generated by N such points is defined to be a random spatial pattern, or equivalently, a realization of a random spatial point process.

---

[1] By module we mean the average distance between a point and its nearest neighbor.

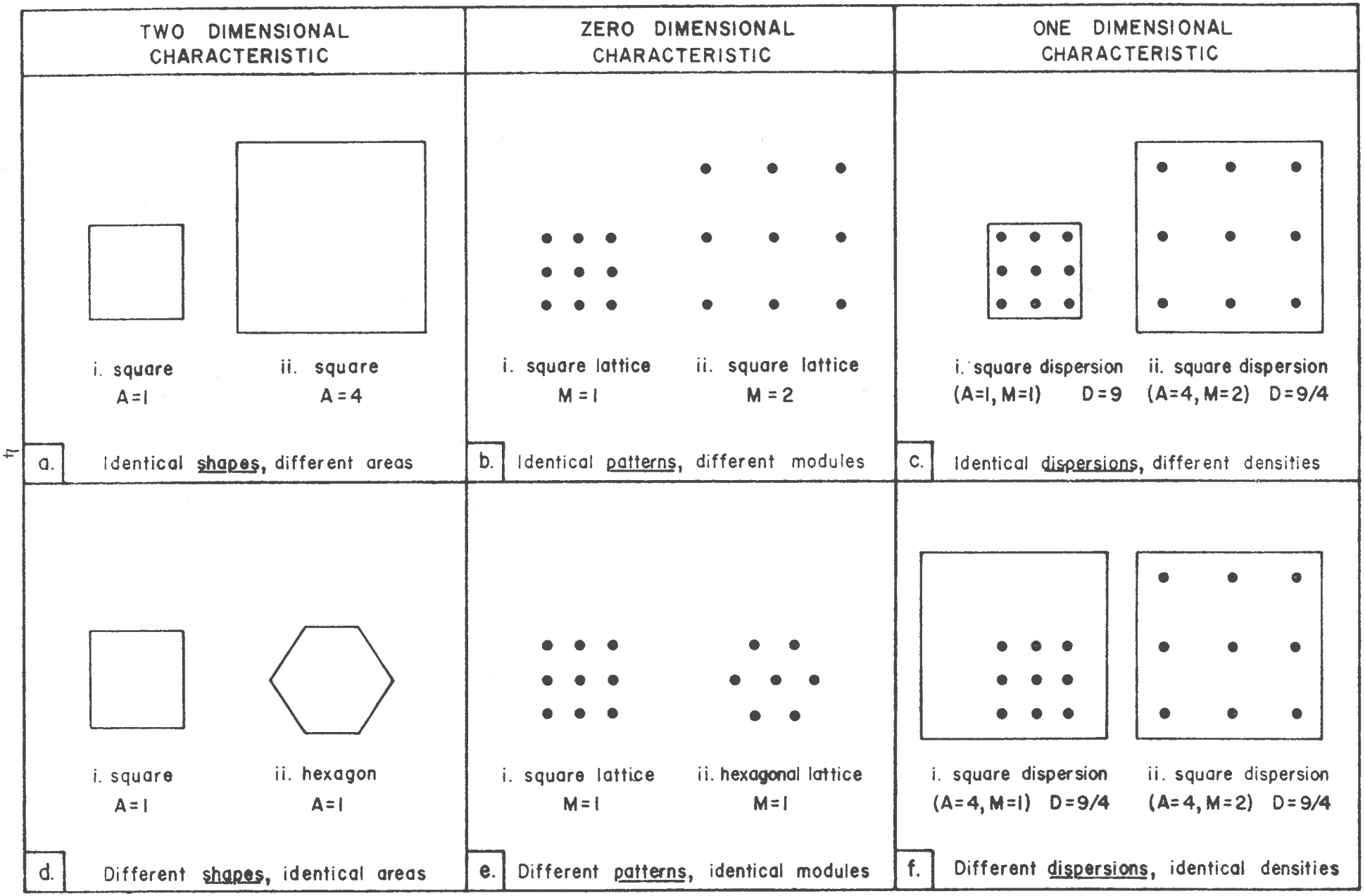


FIG. I.I — CHARACTERISTICS OF SPATIAL ARRANGEMENTS

If N points are located randomly in a planar region, then whether a point falls within a particular subdivision of area A or not can be regarded as an equiprobable event which occurs with probability λA, where λ is the density (number of points per unit area) of this random spatial point pattern. Now consider a square subregion of area a in the plane which is made up of n very small square subdivisions. Assume that these square subdivisions are so small that the probability of more than one point occurring in them is insignificant and tends to zero as n becomes large. Then $A = \frac{a}{n}$, and the probability that a subdivision contains no points is $\left(1 - \frac{\lambda a}{n}\right)$. Since there are $\binom{n}{r}$ ways of combining r subdivisions, each with only a single point, from the n subdivisions, and because each of these combinations has the probability $\left(\frac{\lambda a}{n}\right)^r \left(1 - \frac{\lambda a}{n}\right)^{n-r}$ of occurring, the probability of r points in a square subregion of area a is:

$$P(R = r) = P_r = \lim_{n\to\infty} \binom{n}{r} \left(\frac{\lambda a}{n}\right)^r \left(1 - \frac{\lambda a}{n}\right)^{n-r}$$

$$= \lim_{n\to\infty} \frac{n(n-1)\cdots(n-r+1)}{r!} \frac{(\lambda a)^r}{n^r} \left(1 - \frac{\lambda a}{n}\right)^n \left(1 - \frac{\lambda a}{n}\right)^{-r}$$

$$= \lim_{n\to\infty} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{r-1}{n}\right)\left(1 - \frac{\lambda a}{n}\right)^{-r}\left[\left(1 - \frac{\lambda a}{n}\right)^n \frac{(\lambda a)^r}{r!}\right]$$

And as n becomes infinitely large, the terms outside of the square brackets tend to unity and the terms inside the brackets yield

$$P_r = e^{-\lambda a} \frac{(\lambda a)^r}{r!} \qquad (r = 0, 1, 2, \ldots) \qquad (1.1)$$

The probability distribution defined by (1.1) is known as the Poisson law with the single parameter λa. Its expected value is

$$m_1 = E(r) = \sum_{r=0}^{\infty} re^{-\lambda a} \frac{(\lambda a)^r}{r!} = \lambda a \sum_{r=0}^{\infty} e^{-\lambda a} \frac{(\lambda a)^{r-1}}{(r-1)!}$$

$$= \lambda a \sum_{x=0}^{\infty} e^{-\lambda a} \frac{(\lambda a)^x}{x!} \quad , \quad \text{where } x = r - 1 \ ,$$

$$= \lambda a \ . \tag{1.2}$$

Its second moment is

$$E(r^2) = \sum_{r=0}^{\infty} r^2 e^{-\lambda a} \frac{(\lambda a)^r}{r!} = \sum_{r=0}^{\infty} r(r-1)e^{-\lambda a} \frac{(\lambda a)^r}{r!} + \sum_{r=0}^{\infty} re^{-\lambda a} \frac{(\lambda a)^r}{r!}$$

$$= \lambda^2 a^2 \sum_{r=0}^{\infty} e^{-\lambda a} \frac{(\lambda a)^{r-2}}{(r-2)!} + \lambda a$$

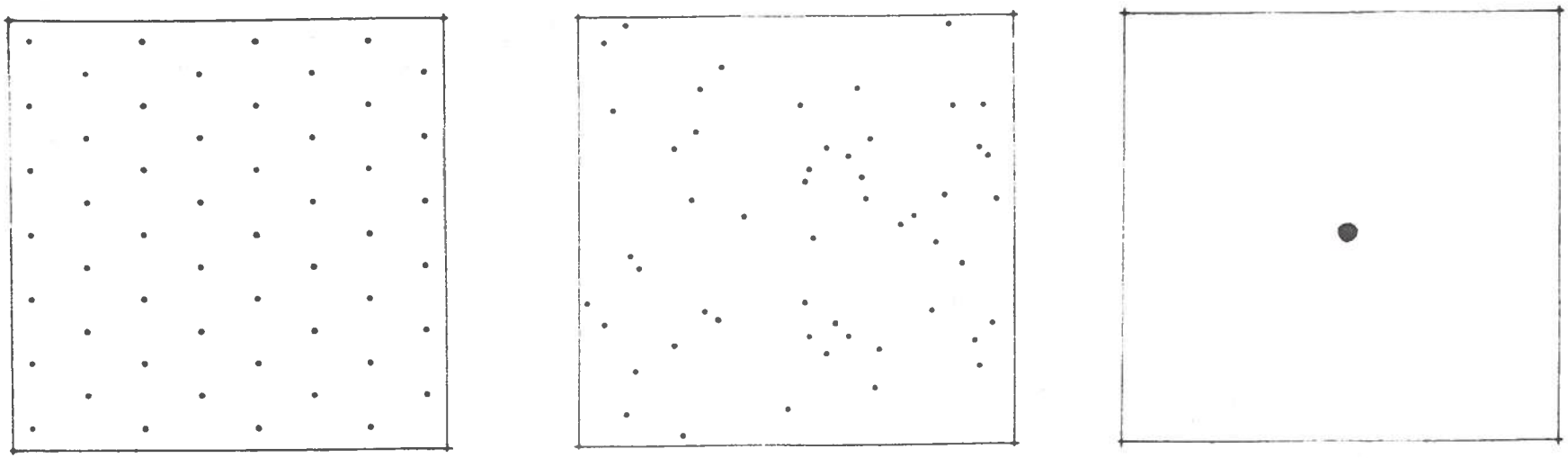
$$= \lambda^2 a^2 + \lambda a \ .$$

Hence, the variance of the Poisson distribution is

$$m_2 = E(r^2) - [E(r)]^2 = \lambda^2 a^2 + \lambda a - \lambda^2 a^2$$

$$= \lambda a \ . \tag{1.3}$$

Figure 1.2b illustrates a random spatial point pattern which contains 52 points. In Figure 1.2 we also present the two extreme point patterns that can be generated with the same 52 points. Figure 1.2a illustrates the case of maximum regularity, or perfect regularity, and Figure 1.2c illustrates the case of maximum clustering, or perfect clustering. The random point pattern lies midway between these two extreme distributions.



a. Perfectly regular point pattern

b. Random point pattern

c. Perfectly clustered point pattern

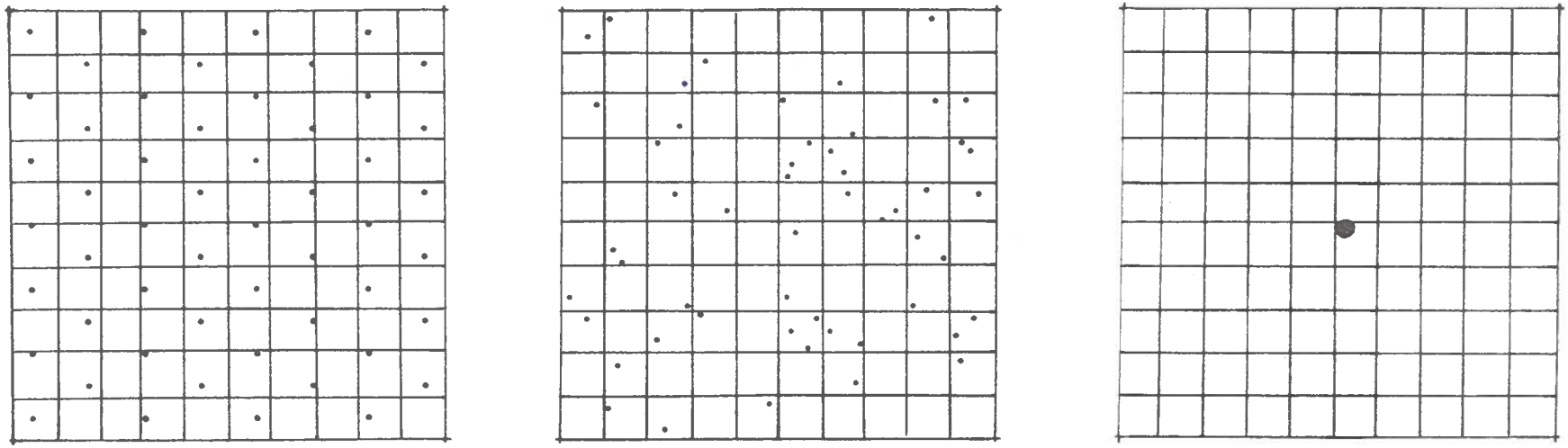
FIG. 1.2--PERFECTLY REGULAR, RANDOM, AND PERFECTLY CLUSTERED SPATIAL POINT PATTERNS (N = 52)

1.3 Quadrat Analysis

Quadrat analysis is the name that has been associated with a class of methods of spatial analysis which have been developed over the past thirty years largely by plant ecologists. In the quadrat method, a planar study region is divided into a grid with cells of equal size, called quadrats, and the number of points in each cell, or in randomly selected cells, is noted. Quadrats generally are square shaped and the analysis strives to obtain indications of the random or nonrandom character of the model that generated the spatial distribution being studied. An observed frequency distribution that does not conform with one expected from a random point process, leads to a rejection of the hypothesis of randomness in favor of an alternative hypothesis of a more regular or a more clustered than random model, depending on the way in which the observed values differ from expected values.

Intuitively, we would expect a regular point process to generate a large number of quadrats with only a single point in them, some empty quadrats, and very few quadrats with more than one point in them. Conversely, we would expect a clustered point-process to produce a very large number of empty quadrats, a few quadrats with one or two points in them, and several quadrats with many points in them. The random point process would be expected to produce results somewhere in between these two extremes.

Figures 1.3a, 1.3b, and 1.3c illustrate the fit of a 10 by 10 grid of quadrats over the three point patterns in Figure 1.2. In the case of the perfectly regular pattern, 52 out of the 100 quadrats have only a single point in them, 48 quadrats are empty, and none has more than a single point. The perfectly clustered pattern, on the other hand, yields 99 empty quadrats and a single quadrat containing all of the 52 points. These results, along with the corresponding histogram for the random point pattern are illustrated in Figure 1.4.

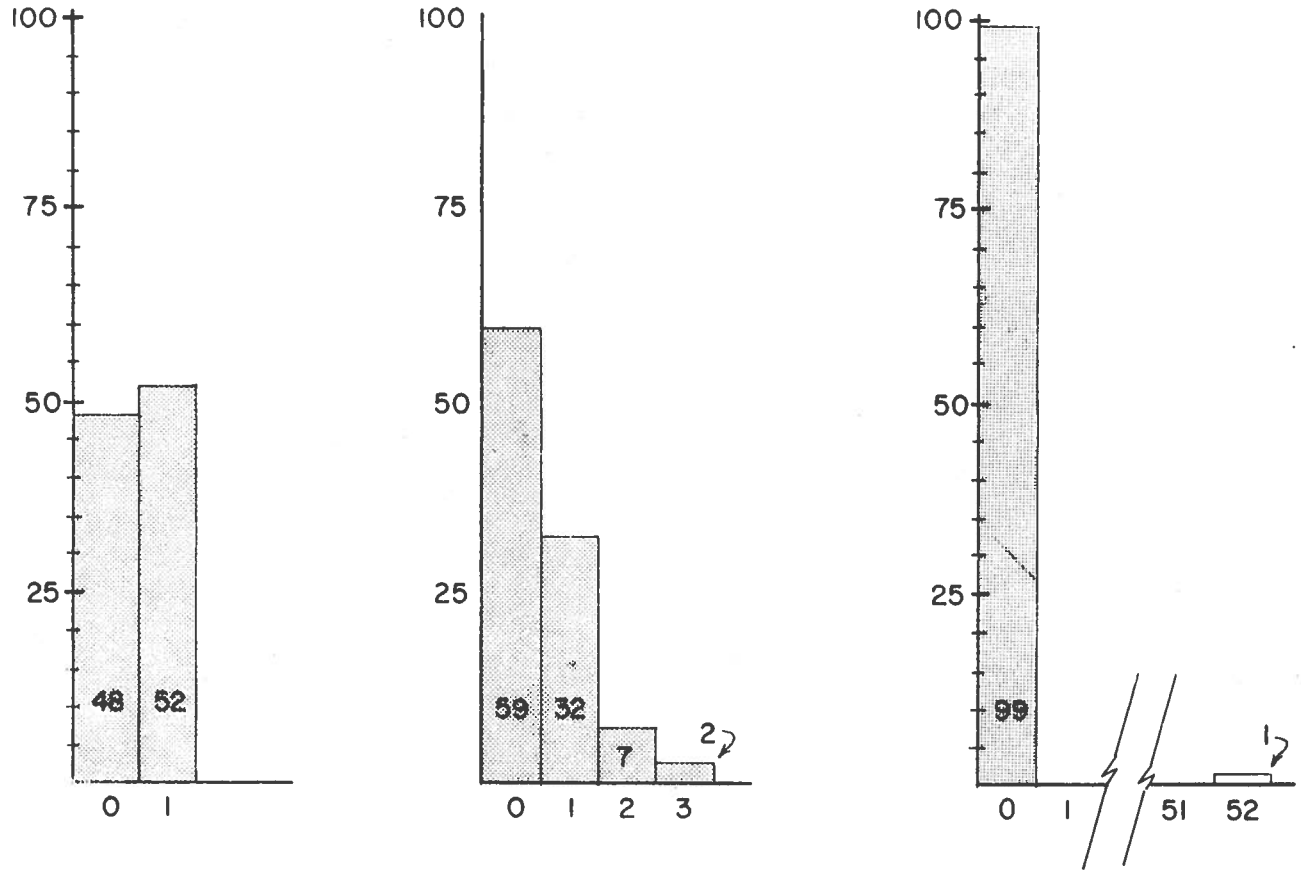


a. Perfectly regular

b. Random

c. Perfectly clustered

FIG. 1.3--QUADRAT COVERAGE OF PERFECTLY REGULAR, RANDOM, AND PERFECTLY CLUSTERED SPATIAL POINT PATTERNS (N = 52)



a. Perfectly regular

b. Random

3. Perfectly clustered

FIG. 1.4--FREQUENCY DISTRIBUTIONS OF PERFECTLY REGULAR, RANDOM, AND PERFECTLY CLUSTERED POINT PATTERNS (N = 52, n = 100)

1.3.1 The Variance-Mean Ratio

The variance of the Poisson distribution is equal to its mean. Hence the ratio of the variance to the mean, for such distributions is equal to unity, and observed point dispersions may be measured for their departure from expected Poisson realizations by testing the significance of the difference between the observed variance-mean ratio and unity. This difference has a standard error of $\sqrt{2/(n-1)}$, where n is the number of observations, and its significance may be statistically tested with a t-test with n - 1 degrees of freedom [Greig-Smith (1964)]. A realization yielding a variance-mean ratio greater than unity indicates a more clustered than random spatial point process, whereas one with a variance-mean ratio less than unity is more likely to result from a more regular than random model.

To illustrate the use of the variance-mean ratio test, we may compute this index for the three point dispersionsin Figure 1.3. We have the following sample statistics:

Perfectly regular	Random	Perfectly clustered
$u_1 = .5200$	$u_1 = .5200$	$u_1 = .5200$
$u_2 = .2521$	$u_2 = .5148$	$u_2 = 27.0400$
$u_2/u_1 = .4848$	$u_2/u_1 = .9899$	$u_2/u_1 = 52.0000$
$t^* = \frac{.4848-1.0}{.1421} = -3.6256$	$t^* = \frac{.9899-1.0}{.1421} = -0.0711$	$t^* = \frac{52.0-1.0}{.1421} = 358.9021$

Each of the above t statistics is distributed as a t distribution with 99 degrees of freedom, and only the random dispersion is not significantly different from the realization that would be expected under a Poisson model, at the 5 per cent level of significance.

1.3.2 The Chi-Square Goodness of Fit Test

An alternative method for measuring the departure of an observed spatial point dispersion from a random one is the chi-square goodness of fit test. This test is described in greater detail in Part III. Thus we merely illustrate its application here.

To apply the chi-square test, we begin by estimating the population mean of the Poisson distribution, $m_1 = \lambda a$, with the sample mean, u_1. Then we compute the expected frequencies:

$$f_r = nP_r = ne^{-u_1}\frac{(u_1)^r}{r!}, \quad (r = 0, 1, 2, \ldots,) \tag{1.4}$$

and measure the departure of observed from expected frequencies by means of the X^2 statistic:

$$X^2 = \sum_{r=0}^{R}\frac{(f_r - nP_r)^2}{nP_r} \tag{1.5}$$

where R denotes the number of frequency classes, f_r denotes the number of observations in each frequency class, n is the sample size $\left(\sum_{r=0}^{R} f_r = n\right)$, and P_r denotes the probability that an observation falls into the rth frequency class, under the hypothesis of random dispersion. Finally, we compare the computed X^2 statistic with tabulated percentiles for a chi-square distribution with R - 1 degrees of freedom. Values of X^2 that exceed the critical value obtained from the chi-square tables lead to a rejection of the null hypothesis of randomness.

Table 1.1 illustrates the application of the chi-square goodness of fit test to the three spatial point dispersions in Figure 1.3. Note that in the case of the perfectly regular point pattern we have no degrees of

Table 1.1 Fitting the Poisson Distribution to a Perfectly Regular, a Random, and a Perfectly Clustered Spatial Point Pattern.

(N = 52, n = 100)

Number of points per quadrat	Observed Frequencies			Expected frequency with Poisson model
	Perfectly regular	Random	Perfectly clustered	
0	48	59	99	59.45
1	52	32	0	30.92
2		7	0	8.04
3+		2	1	1.59
N =	100	100	100	100.00
X^2 =	--	0.60	64.96	
P(.05) =	--	3.84	3.84	

freedom left to carry out the chi-square test, and observe that the perfectly clustered point pattern leads to an unqualified rejection of the null hypothesis of a random point process.

1.3.3 Some Difficulties with the Quadrat Method

A host of theoretical and practical problems are associated with the quadrat method. Some of these will be considered in detail in this paper. For example, it is clear that the size of the quadrat influences the findings. This problem is discussed in Section 3.3.4. Furthermore, the method is density dependent. Hence it is more appropriately used as a technique for analyzing dispersion rather than pattern. Another difficulty is that the placement of the grid system over the point pattern also can influence the results.

Finally, there are a number of vexing problems of inference associated with the fitting of non-random distributions. For example, several different probability models can give rise to the same expected frequency distribution, a point which will be illustrated, in Part II, with the negative binomial model. And the chi-square goodness of fit test requires that observations be pooled in way that provides an expectation of 5 units for each frequency category. We shall see, in Section 3.3.3, that this can lead to serious difficulties.

1.4 Nearest Neighbor Analysis

A method for gauging the departure from randomness of an observed spatial point pattern that does not suffer from some of the faults of the quadrat method is the nearest neighbor technique. This technique utilizes the distribution, in a random point pattern, of the distance between a point and its closest neighboring point. The derivation of this particular distribution, which appears below, is adapted from the argument in Clark and Evans (1954).

1.4.1 The Distribution of the Nearest Neighbor Distance in a Random Point Pattern

Consider the distance of any point to its nearest neighbor in a random pattern. First, let us center a circular quadrat of area a over the point. The probability of finding r points in this quadrat is given by the Poisson distribution

$$P_r = e^{-\lambda a} \frac{(\lambda a)^r}{r!} .$$

Hence, the probability of finding no points is

$$P_0 = e^{-\lambda a} \quad ,$$

or, if the radius of the circular quadrat is d, then

$$P_0 = e^{-\lambda \pi d^2} \tag{1.6}$$

But this also is the probability that the distance to the nearest neighboring point, D say, is greater than d. Thus

$$P(D > d) = P_0 = e^{-\lambda \pi d^2} \quad .$$

Consequently,

$$P(D \leq d) = 1 - P_0 = 1 - e^{-\lambda \pi d^2} = G(d) \quad , \tag{1.7}$$

where G(d) is the cumulative distribution function of the nearest neighbor distance, d.

Differentiating (1.7), we obtain the associated probability density function

$$g(d) = G'(d) = 2\lambda \pi d e^{-\lambda \pi d^2} \quad , \tag{1.8}$$

and conclude, therefore, that the square of the nearest neighbor distance is distributed as a negative exponential distribution with mean $\frac{1}{\lambda \pi}$ and variance $\frac{1}{\lambda^2 \pi^2}$.[2] Alternatively, we may work directly with (1.8) to obtain the mean and variance of that distribution. We have, then, that

---

[2] Let $t = d^2$, then $\frac{dt}{dd} = 2d$ and $G(t) = \lambda \pi e^{-\lambda \pi t}$.

$$E(d) = \frac{1}{2\sqrt{\lambda}} \tag{1.9}$$

and

$$V(d) = \frac{4 - \pi}{4\lambda\pi} \quad . \tag{1.10}$$

1.4.2 Sample Statistics

Since in a random point pattern, each nearest neighbor distance is distributed independently from all others and follows the distribution defined by (1.8), we may conclude that the average nearest neighbor distance

$$S = \frac{\sum_{i=1}^{N} d_i}{N} = \bar{d}_O \quad \text{say,} \tag{1.11}$$

is distributed approximately as the normal distribution with mean $\bar{d}_E = \frac{1}{2\sqrt{\lambda}}$ and variance $V(d) = \frac{4 - \pi}{4\lambda\pi N}$. Thus

$$\phi = \frac{\bar{d}_O - \bar{d}_E}{\sigma_{\bar{d}_E}} \quad , \tag{1.12}$$

where

$$\sigma_{\bar{d}_E} = \frac{.26136}{\sqrt{\lambda N}} \tag{1.13}$$

approximately follows a normal distribution with zero mean and unit variance.

Clark and Evans (1954) suggest the use of the index

$$D = \frac{\bar{d}_O}{\bar{d}_E} \tag{1.14}$$

to compare different point patterns. They show that a perfectly regular point pattern leads to a D = 2.1491, a random point pattern to a D = 1

and, of course, a perfectly clustered pattern to a D = 0.

To illustrate the application of the nearest neighbor test, we may compute this index for the three point patterns in Figure 1.2. This yields an estimate of λ that is equal to .0052 and the following sample statistics:[3]

Perfectly regular	Random	Perfectly clustered
$\bar{d}_0 = 15.00$	$\bar{d}_0 = 7.59$	$\bar{d}_0 = 0.00$
D = 2.16	D = 1.09	D = 0.00
$\Phi = \frac{15.00 - 6.93}{.52} = 15.52$	$\Phi = \frac{7.59 - 6.93}{.52} = 1.27$	$\Phi = \frac{0.00 - 6.93}{.52} = -13.32$

Once again, only the random dispersion is not significantly different from the realization that would be expected under a Poisson model, at the 5 per cent level of significance.

1.4.3 Some Difficulties with the Nearest Neighbor Method

Although free of some of the difficulties which were associated with the quadrat method (such as the problem of quadrat size), nearest neighbor analysis also suffers from several serious problems. Some problems of quadrat methods, such as density dependence, are shared by the nearest neighbor technique. Others are unique to the method itself. For example, sole reliance on the nearest neighbor distance may lead to erroneous conclusions if the point pattern has a configuration of "clumps," such as the point pattern in Figure 1.5, below, which yields a nearest neighbor distance of zero, thereby indicating a perfect clustering when it does not exist. To deal with this problem, Thompson (1956) and others have suggested the use of subareas or n^{th} nearest neighbors. For example, the average distance to the first, second, third, fourth, and fifth nearest neighbors in the spatial point patterns in Figure 1.2 are:

---

[3] In computing the nearest neighbor statistics we measured distances between points to the nearest tenth of a quadrat edge. Hence the estimate of λ used here is .0052 or $\frac{1}{10^2}$ of the corresponding estimate of λ used in Section 1.3.2.

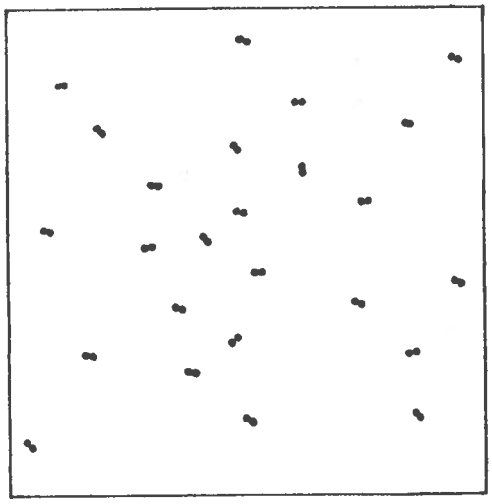


FIG. 1.5--A PATTERN OF TWO-POINT CLUMPS (N = 52)

	Perfectly regular	Random	Perfectly clustered
First:	15.00	7.59	0.00
Second:	15.00	11.49	0.00
Third:	15.43	14.16	0.00
Fourth:	17.12	16.57	0.00
Fifth:	19.17	18.87	0.00

Perhaps the most serious limitation of nearest neighbor analysis lies in the difficulty of obtaining the probability distributions of nearest neighbor distances in spatial point processes other than the random one. Thus, although non-random patterns may be ranked according to their degree of regularity or clustering, we cannot infer much about the probability model that generated them. This is the principal reason why we have adopted the quadrat method as the technique for analyzing urban dispersion in this paper.

PART II

THEORETICAL DISTRIBUTIONS IN QUADRAT ANALYSIS

2.1 The Fundamental Component Distributions

In order to determine whether the dispersion of a particular set of points relative to a pre-defined study region is regular, random or clustered, we must first define in mathematical terms of what distributional form such dispersions are likely to exhibit. The normal procedure is to adopt the Poisson distribution as the mathematical definition of "randomness." We shall now rederive this distribution in a way that will allow us to extend our analysis to non-random dispersions.

2.1.1 Random, Regular and Clustered Spatial Dispersion

Imagine a study region that has been gridded into square cells of a given dimension. Assume that initially (i.e., at time $t = 0$), none of the cells contains any points, and let $p(r, t)$ be the probability that an individual grid cell has r points by time t. Assume that during the time interval $(t, t + dt)$ a point locates in a particular cell, which already has r points, with probability $f(r, t)dt$ and that the time interval is short enough so that no more than one point may locate in a given cell during a single time interval. It then follows that:

$$p(0, t + dt) = p(0, t)[1 - f(0, t)dt]$$

$$p(r, t + dt) = p(r, t)[1 - f(r, dt)dt] + p(r - 1, t)f(r - 1, t)dt \quad , \qquad (r = 1, 2, 3, \ldots)$$

and, subtracting $p(r, t)$ from both sides of the above equations, dividing by dt, and taking the limit as $dt \to 0$, we have:

$$\frac{dp}{dt}(0, t) = - f(0, t)p(0, t)$$

$$\frac{dp}{dt}(r, t) = - f(r, t)p(r, t) + f(r - 1, t)p(r - 1, t) \quad (r = 1, 2, 3, \ldots) \tag{2.1}$$

Let us multiply the first equation in (2.1) by unity, the second by s, the third by s^2, and, in general, the n^{th} by n - 1. Summing the resulting equations, we obtain

$$\frac{d}{dt}\left[\sum_{r=0}^{\infty} p(r, t)s^r\right] = (s - 1)\left[\sum_{r=0}^{\infty} f(r, t)p(r, t)s^r\right] ,$$

or more compactly,

$$\frac{d}{dt} F(s) = (s - 1)H(r, s) \quad , \tag{2.2}$$

where F(s) is the probability generating function of the random variable r.[4] To find F(s) we need to solve the single differential equation in (2.2). Depending on the assumption we make concerning f(r, t), our resulting distribution will be either a <u>random</u> one, a <u>regular</u> one, or a <u>clustered</u> one. Note that f(r, t) is a probability that is contingent on the value of r. It expresses the probability that a cell with r points will obtain yet another one in the time interval (t, t + dt). If this probability is independent of the number of points already in the cell, we shall refer to the spatial distribution as <u>random</u> dispersion. If, on the other hand, this probability declines as the number of existing points in the cell increases, then we define the spatial dispersion to be a <u>regular</u> one. Finally, if the probability increases as the number of existing points in the cell increases, we shall refer to the resulting spatial distribution as <u>clustered</u> dispersion.

---

[4] For a full discussion of the role and use of probability generating functions in discrete probability theory see Feller (1957), pp. 248-261.

In the following sections we shall be more specific and assume functional values for f(r, t) and derive the distributional consequences of these assumptions. The resulting distributions will be called the fundamental component distributions of quadrat analysis, because most of the other more complex distributions, that are commonly used in quadrat analysis, can be derived as particular mixtures or combinations of these basic distributions.

2.1.2 Random Spatial Dispersion: The Poisson Distribution

Assume that the probability that a cell receives a point during the time interval (t, t + dt) is completely independent of the number of such points already in the cell. Then

$$f(r, t) = f(t)$$

$$H(r, s) = \sum_{r=0}^{\infty} f(t)p(r, t)s^r = f(t)F(s) \qquad (2.3)$$

and (2.2) becomes

$$\frac{d}{dt} F(s) = (s - 1)f(t)F(s) \quad ,$$

with the solution

$$F(s) = e^{-\lambda a(1-s)} \quad ,$$

where

$$\lambda a = \int_0^t f(t)dt \quad . \qquad (2.4)$$

Equation (2.4) is easily recognizable as the probability generating function (p.g.f.) of the Poisson distribution with parameter λa. Hence, for any point in time, $\bar{t}$, say:

$$p(r, \bar{t}) = P_r = e^{-\lambda a} \frac{(\lambda a)^r}{r!} \quad (r = 0, 1, 2, \ldots) \qquad (2.5)$$

As a check, let us compute the p.g.f. of the above Poisson distribution.

By definition,

$$F(s) = \sum_{r=0}^{\infty} P_r s^r \quad .$$

Hence

$$F(s) = \sum_{r=0}^{\infty} e^{-\lambda a} \frac{(\lambda a)^r}{r!} s^r = e^{-\lambda a} \sum_{r=0}^{\infty} \frac{(\lambda as)^r}{r!} \quad ,$$

and, introducing $e^{-\lambda as}$ and $e^{\lambda as}$ ($e^{-\lambda as} \cdot e^{\lambda as} = 1$), we have:

$$F(s) = e^{-\lambda a} \cdot e^{\lambda as} \sum_{r=0}^{\infty} e^{-\lambda as} \frac{(\lambda as)^r}{r!}$$

$$= e^{-\lambda a(1-s)}(1)$$

$$= e^{-\lambda a(1-s)} \quad .$$

Using the well known relationships

$$m_1 = F'(1)$$

and

$$m_2 = F''(1) + F'(1) - [F'(1)]^2 \quad ,$$

we may derive the mean and variance of the Poisson distribution. Thus

$$m_1 = F'(1) = \lambda a$$

$$m_2 = F''(1) + F'(1) - [F'(1)]^2 = \lambda a \quad .$$

Note, once again, that the ratio of the variance to the mean of this distribution is always unity, i.e., $\frac{m_2}{m_1} = \frac{\lambda a}{\lambda a} = 1$.

2.1.3 Regular Spatial Dispersion: The Binomial Distribution

For our next derivation, assume that the probability that a point locates in a cell varies inversely with the number of points already in that cell. More specifically, let c/b be an integer and

$$f(r, t) = c - br \quad \text{for } br < c > 0$$
$$= 0 \qquad \text{otherwise} \tag{2.6}$$

Then

$$H(r, s) = \sum_{r=0}^{\infty} (c - br)p(r, s)^r = c \sum_{r=0}^{\infty} p(r, s)s^r - b \sum_{r=0}^{\infty} rp(r, s)s^r$$

$$= cF(s) - bs \frac{\partial}{\partial s} F(s) \quad ,$$

and (2.2) becomes

$$\frac{\partial}{\partial t} F(s) = (s - 1)\left[cF(s) - bs \frac{\partial}{\partial s} F(s)\right] \quad , \tag{2.7}$$

with the solution

$$F(s) = [e^{bt} + (e^{bt} - 1)s]^{c/b} = (1 - p + ps)^n \quad , \tag{2.8}$$

where $p = e^{bt} - 1$ and $n = c/b$.

Equation (2.8) is the probability generating function of the binomial distribution

$$P_r = \binom{n}{r} p^r(1 - p)^{n-r} \quad (r = 0, 1, 2, \ldots, n) \quad . \tag{2.9}$$

As a check, we may compute the p.g.f. of the above binomial distribution, as follows:

$$F(s) = \sum_{r=0}^{\infty} p_r s^r = \sum_{r=0}^{n} \binom{n}{r} p^r (1 - p)^{n-r} s^r$$

$$= \sum_{r=0}^{n} \binom{n}{r} (ps)^r (1 - p)^{n-r}$$

$$= (1 - p + ps)^n \quad .$$

Differentiating to derive the mean and variance, we obtain

$$m_1 = F'(1) = np$$

and

$$m_2 = F''(1) + F'(1) - [F'(1)]^2 = np(1 - p) \quad .$$

Observe that the variance-mean ratio of this distribution is $1 - p$ which is always less than unity.

When n is large and p is small, i.e., $n \to \infty$ and $p \to 0$, such that $np = \lambda a$, the Poisson distribution provides a reasonably close approximation to the binomial distribution. We may prove this by means of a limiting argument on the probability generating function of the binomial distribution. For the binomial distribution,

$$F(s) = (1 - p + ps)^n \quad .$$

If we let $n \to \infty$ and $p \to 0$, such that $np = \lambda a$ remains fixed, then

$$(1 - p + ps)^n = \left[1 - \frac{\lambda a(1 - s)}{n}\right]^n \quad ,$$

and

$$\lim_{\substack{n\to\infty \\ p\to 0}} \left[1 - \frac{\lambda a(1 - s)}{n}\right]^n = e^{-\lambda a(1-s)} \quad ,$$

which we have already shown is the probability generating function of the Poisson distribution.

2.1.4 Clustered Spatial Dispersion: The Negative Binomial Distribution

For our final derivation, assume that the probability that a point locates in a cell is directly related to the number of points already in that cell. In particular, assume that c/b is an integer and

$$f(r, t) = c + br \quad (c > 0, b > 0) \quad . \tag{2.10}$$

Then

$$H(r, s) = \sum_{r=0}^{\infty} (c + br)p(r, s)s^r = c \sum_{r=0}^{\infty} p(r, s)s^r + b \sum_{r=0}^{\infty} rp(r, s)s^r$$

$$= cF(s) + bs \frac{\partial}{\partial s} F(s) \quad ,$$

and (2.2) becomes

$$\frac{\partial}{\partial t} F(s) = (s - 1)\left[cF(s) + bs \frac{\partial}{\partial s} F(s)\right] \quad . \tag{2.11}$$

Equation (2.11) is a simple partial differential equation with the solution

$$F(s) = [e^{bt} - (e^{bt} - 1)s]^{-c/b} = (1 + p - ps)^{-k} \tag{2.12}$$

where $p = e^{bt} - 1$ and $k = c/b$.

Equation (2.12) is the probability generating function of the negative binomial distribution[5]

---

[5] If k = c/b is not an integer, then we may use the gamma function to express the negative binomial distribution, as follows:

$$P_r = \frac{\Gamma(k + r)}{r!\Gamma(k)} \left[\frac{p}{1 + p}\right]^r \left[\frac{1}{1 + p}\right]^k \quad .$$

$$P_r = \binom{k + r - 1}{r} \left[\frac{p}{1 + p}\right]^r \left[\frac{1}{1 + p}\right]^k \qquad (2.13)$$

We may compute the p.g.f. of the above distribution to find:

$$F(s) = \sum_{r=0}^{\infty} P_r s^r = \sum_{r=0}^{\infty} \binom{k + r - 1}{r} \left[\frac{p}{1 + p}\right]^r \left[\frac{1}{1 + p}\right]^k s^r$$

$$= \left[\frac{1}{1 + p}\right]^k \sum_{r=0}^{\infty} \binom{k + r - 1}{r} \left[\left(\frac{p}{1 + p}\right) s\right]^r$$

$$= \left[\frac{1}{1 + p}\right]^k \left[1 - \left(\frac{p}{1 + p}\right) s\right]^{-k}$$

$$= (1 + p - ps)^{-k} \quad .$$

Differentiating F(s) to derive the mean and variance, we find

$$m_1 = F'(1) = kp$$

and

$$m_2 = F''(1) + F'(1) - [F'(1)]^2 = kp(1 + p) \quad .$$

Note that the variance-mean ratio for this distribution is $1 + p$, which is always greater than unity.

When k is large and p is small, i.e., $k \to \infty$ and $p \to 0$, such that $kp = \lambda a$, the negative binomial distribution approaches the Poisson distribution. As with the binomial case, this can be established by an argument using probability generating functions. For the negative binomial distribution,

$$F(s) = (1 + p - ps)^{-k} \quad .$$

Hence, if we let $k \to \infty$ and $p \to 0$, such that $kp = \lambda a$ remains fixed, then

$$(1 + p - ps)^{-k} = \left(1 + \frac{\lambda a}{k} - \frac{\lambda as}{k}\right)^{-k}$$

and

$$\lim_{\substack{k\to\infty \\ p\to 0}} \left[\frac{1}{1 + \frac{\lambda a(1 - s)}{k}}\right]^{k} = \frac{1}{e^{\lambda a(1-s)}}$$

$$= e^{-\lambda a(1-s)}$$

which is the p.g.f. of the Poisson distribution.

Occasionally it is more convenient to operate with the following alternative forms of the negative binomial distribution:

$$P_r = \binom{k + r - 1}{r} \left(\frac{m}{m + k}\right)^{r} \left(\frac{k}{m + k}\right)^{k} \tag{2.14}$$

$$= \binom{k + r - 1}{r} p^{r} q^{-k-r} \qquad (q = 1 + p) \tag{2.15}$$

$$= \binom{k + r - 1}{r} Q^{r} P^{k} \qquad (Q = 1 - P = p) \quad , \tag{2.16}$$

and their respective probability generating functions:

$$F(s) = \left(1 + \frac{m}{k} - \frac{ms}{k}\right)^{-k} \tag{2.17}$$

$$= (1 + p - ps)^{-k} \tag{2.18}$$

$$= \left(\frac{P}{1 - Qs}\right)^{k} \quad . \tag{2.19}$$

2.1.5 Simulation of the Fundamental Component Distributions[6]

Frequently it is desirable to obtain theoretical spatial point patterns which have specific distributional properties. Such simulated patterns may be used to test the influence of alternative quadrat sizes, for example. And they provide the synthetically created data that are so useful in the study of the efficiency of alternative parameter estimation methods. The above derivations of the Poisson, binomial and negative binomial distributions suggest a way in which such component distributions may be simulated.

Recall the following three fundamental stochastic processes defined by (2.3), (2.6) and (2.10), respectively:

Poisson: $f(r, t) = f(t)$;

Binomial: $f(r, t) = c - br$, for $br < c > 0$,

$= 0$ otherwise

Negative Binomial: $f(r, t) = c + br$, for $b > 0$ and $c > 0$.

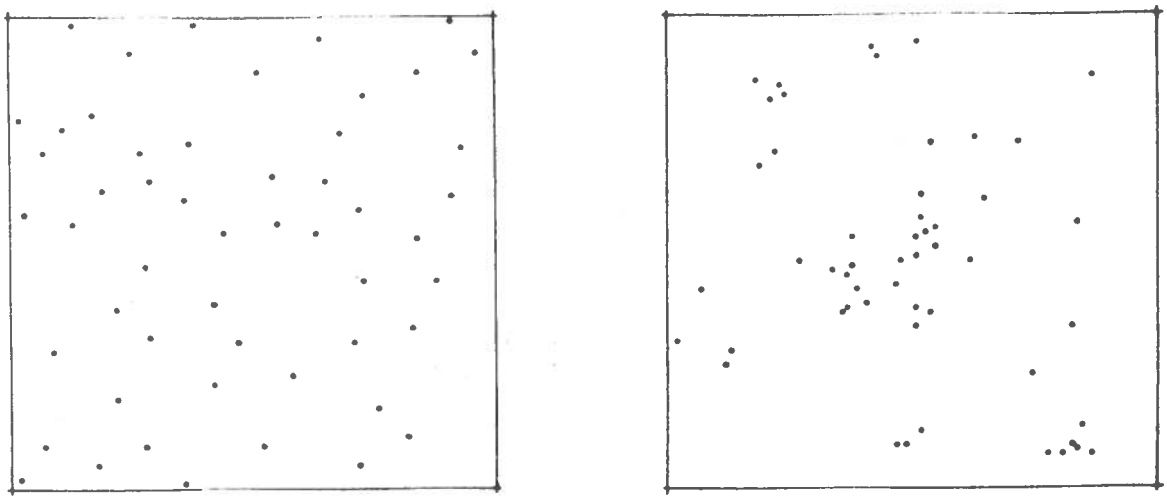
If we locate N points in a grid according to one of these stochastic processes the resulting point dispersion should approximate a realization of the stochastic process that we have selected. Thus, for example, if in a grid of n squares we locate N points according to the process defined by (2.3), the resulting point dispersion should be an approximate realization of the Poisson distribution with mean $m_1 = \frac{N}{n}$. Alternatively, if each of the N points is located, sequentially, according to (2.10), the resulting point dispersion should be an approximate realization of the negative binomial

---

[6] For a more detailed exposition of computer generated spatial patterns, see: Martin, Gomar and Rogers (1969).

distribution with parameters that are functions of the constants c and b.

The above procedures yield point dispersions, i.e., number of points in each quadrat. However, they do not provide point patterns. To obtain the Poisson, binomial and negative binomial patterns illustrated in Figures 1.2 and 2.1, therefore, we proceeded as follows. First, we located 52 points into 100 quadrats according to the rules specified by (2.3), (2.6) and (2.10). Then, within each quadrat, we located the requisite number of points randomly. The resulting realizations seem to be adequate representatives of their respective asymptotic distributions, according to Table 2.1.



a. A regular (binomial) point pattern

b. A clustered (negative binomial) point pattern

FIG. 2.1--A REGULAR (BINOMIAL) AND A CLUSTERED (NEGATIVE BINOMIAL) SPATIAL POINT PATTERN (N = 52)

TABLE 2.1--OBSERVED AND EXPECTED DISTRIBUTIONS OF THE SIMULATED PATTERNS

No. of Points per Cell	Perfectly Regular[a]			Regular[a]			Random[a]				Clustered			Perfectly Clustered			
	Observed Frequency	B. n = 1 p = 0.5250	P. m = 0.5200	Observed Frequency	B. n = 2 p = 0.2600	P. m = 0.5200	Observed Frequency	B. n = 3 p = 0.1733	P. m = 0.5200	Observed Frequency	B. n = 5 p = 0.1040	P. m = 0.5200	N.B. m = 0.5200 k = 0.6781	Observed Frequency	B. n = 52 p = 0.0100	P. m = 0.5200	N.B. m = 0.5200 k = 0.0102
0	48	48.00	59.45	54	54.76	59.45	59	56.49	59.45	67	57.75	59.45	67.98	99	59.30	59.45	96.05
1	52	52.00	40.55	40	38.48	30.92	32	35.54	30.92	23	33.51	30.92	20.01	0	31.15	30.92	0.96
2				6	6.76	9.63	7	7.45	8.04	5	7.78	8.04	7.29	0	8.02	8.04	0.48
3							2	0.52	1.59	2	0.90	1.39	2.82	0	1.35	1.39	0.31
4										2	0.05	0.18	1.13	0	0.17	0.18	0.23
5										1		0.02	0.78	0	0.02	0.02	0.18
6														0			0.15
7														0			0.13
8														0			0.11
9														0			0.09
10+														1			1.31
X^2		--	--		0.16	4.54		0.60	0.08		4.96	3.00	[2.05][b]		65.39	64.96	--
$P_{.05}$		--	--		3.84	3.84		3.84	3.84		3.84	3.84	5.99		3.84	3.84	
u_1 = 0.5200				u_1 = 0.5200			u_1 = 0.5200			u_1 = 0.5200				u_1 = 0.5200			
u_2 = 0.2521		D = 2.16		u_2 = 0.3733		D = 1.44	u_2 = 0.5148		D = 1.09	u_2 = 0.9188		D = 0.75		u_2 = 27.0400		D = 0.00	
u_2/u_1 = 0.4848				u_2/u_1 = 0.7180			u_2/u_1 = 0.9899			u_2/u_1 = 1.7669				u_2/u_1 = 52.0000			

[a] Moment estimators do not exist for the Negative Binomial distribution because the sample variance is smaller than the sample mean.

[b] $[X^2] = X^2$ statistic computed with grouping ≥ 1 instead of ≥ 5.

2.2 Compound and Generalized Distributions

The above three fundamental component distributions can be combined in different ways to form new distributions. We may, for example, "compound" one distribution with another. Or one component distribution may be used to "generalize" another.

2.2.1 Definitions and Notation

Following Gurland (1957), we shall adopt the following definitions and notation for compound and generalized distributions:

Compound Distribution. If X_1 is a random variable with density function $f_1(x_1|\theta)$ and moment generating function (m.g.f.) $M_{x_1}(t)$ for a given θ, and if θ is regarded as a random variable X_2 with density function $f_2(x_2)$ and m.g.f. $M_{x_2}(t)$, then the random variable $R \equiv X_1 \wedge X_2$ with density function

$$P_r = P(R = r) = \int_{-\infty}^{\infty} f_1(r|cx_2)f_2(x_2)dx_2 \quad , \tag{2.20}$$

and m.g.f. $M_r(t)$ is called the compound X_1 variable with respect to the compounder X_2. Here, c is an arbitrary constant.

Generalized Distribution. If X_1 is a random variable with density function $f_1(x_1)$ and p.g.f. $F_1(s)$, and if X_2 is a random variable with density function $f_2(x_2)$ and p.g.f. $F_2(s)$, then the random variable

$$R \equiv X_1 \vee X_2 = \sum_{i=1}^{X_1} X_{2i} \tag{2.21}$$

with density function

$$P_r = \sum_{x_1=0}^{\infty} P(R = r|X_1 = x_1)P(X_1 = x_1) \tag{2.22}$$

and p.g.f. $F_3(s)$, is called the generalized X_1 variable with respect to the generalizer X_2, where $X_{2i}(i = 1, 2, \ldots)$ is a family of independent and

identically distributed integer-valued random variables which are independent of X_1.

Gurland (1957) shows that if X_1 has a p.g.f. of the form $[F_3(s)]^{\theta}$, where θ is the given parameter, then the distributions of $X_1 \wedge X_2$ and $X_1 \vee X_2$ are equivalent. In particular, as we shall show below,

$$\text{Poisson} \wedge \text{Poisson} \sim \text{Poisson} \vee \text{Poisson} \tag{2.23}$$

and

$$\text{Poisson} \wedge \text{Gamma} \sim \text{Poisson} \vee \text{Logarithmic} \tag{2.24}$$

The former yields the Neyman Type A distribution; the second produces the negative binomial distribution.

For the compound distribution in (2.20) we can establish the following useful relationship involving moment generating functions:[7] Let $M_x(\theta)$ denote the moment generating function of the random variable X, and let f(x) denote its probability density function. Then, by definition,

$$M_x(\theta) = E(e^{\theta x}) = \int_{-\infty}^{\infty} e^{\theta x} f(x)dx \quad , \tag{2.25}$$

and the random variable $R \equiv X_1 \wedge X_2$ with density function (2.20) has the moment generating function

$$M_r(\theta) = \int_{-\infty}^{\infty} e^{\theta r} P_r dr = \int_{-\infty}^{\infty} e^{\theta r} \left[\int_{-\infty}^{\infty} f_1(r|cx_2) f_2(x_2) dx_2 \right] dr$$

---

[7] Moment generating functions play a role in continuous probability theory that is analogous to the role played by probability generating functions in discrete probability theory. However, whereas the former can be used with both classes of probability distributions, the latter can be applied only to discrete probability distributions. See: Parzen (1960), pp. 215-223.

$$M_r(\theta) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} e^{\theta r} f_1(r|cx_2)dr \right] f_2(x_2)dx_2$$

$$= \int_{-\infty}^{\infty} \left[M_{r|cx_2}(\theta) \right] f_2(x_2)dx_2$$

$$= \int_{-\infty}^{\infty} E(e^{\theta r}|cX = cx_2) f_2(x_2)dx_2 \quad . \tag{2.26}$$

Similarly, for the generalized distribution in (2.22) we can establish the following useful relationship involving probability generating functions:

$$P_r = P(R = r) = \sum_{x_1=0}^{\infty} P(R = r|X_1 = x_1)P(X_1 = x_1)$$

and

$$F_3(s) = \sum_{r=0}^{\infty} P_r s^r = \sum_{r=0}^{\infty} s^r \sum_{x_1=0}^{\infty} P(R = r|X_1 = x_1)P(X_1 = x_1)$$

$$= \sum_{x_1=0}^{\infty} P(X_1 = x_1) \sum_{r=0}^{\infty} P(R = r|X_1 = x_1)s^r$$

But, for fixed x_1, the distribution of R is the x_1-fold convolution of $\{P(X_{2i} = j)\}$, i.e., $\{P(X_{2i} = j)\}^{x_1 *}$. Hence, the p.g.f. of $P(R = r|X_1 = x_1) = [F_2(s)]^{x_1}$. Consequently

$$F_3(s) = \sum_{x_1=0}^{\infty} P(X_1 = x_1)[F_2(s)]^{x_1}$$

$$= F_1[F_2(s)] \quad . \tag{2.27}$$

2.2.2 Compound Poisson Distributions

Let X_1 be a Poisson distributed random variable with mean λa. Then the random variable $R \equiv X_1 \wedge X_2$ with density function (2.20) is called the compound Poisson variable and its distribution is referred to as a compound Poisson distribution. Whence

$$M_{r|cx_2}(\theta)E(e^{\theta r}|CX_2 = cx_2) = E(e^{\theta r}|A\Lambda = a\lambda) = \sum_{r=0}^{\infty} e^{\theta r} e^{-\lambda a} \frac{(\lambda a)^r}{r!}$$

$$= e^{-\lambda a} e^{\lambda a e^{\theta}} \sum_{r=0}^{\infty} e^{-\lambda a e^{\theta}} \frac{(\lambda a e^{\theta})^r}{r!}$$

$$= e^{-\lambda a(1-e^{\theta})} , \qquad (2.28)$$

and

$$M_r(\theta) = \int_{-\infty}^{\infty} \left[M_{r|cx_2}(\theta) \right] f_2(x_2)dx_2$$

$$= \int_{-\infty}^{\infty} e^{-cx_2(1-e^{\theta})} f_2(x_2)dx_2$$

$$= M_{cx_2}(e^{\theta} - 1) = M_{x_2}[c(e^{\theta} - 1)] = M_{\lambda}[a(e^{\theta} - 1)] \qquad (2.29)$$

Neyman's Type A Distribution. One of the simplest and most common compound distributions is obtained by introducing another Poisson variable as the compounder, X_2. Then

$$f_2(x_2) = e^{-m} \frac{m^{x_2}}{x_2!} = e^{-m} \frac{m^{\lambda}}{\lambda!} ,$$

and, using (2.29),

$$M_r(\theta) = e^{-m[1-e^{a(e^{\theta}-1)}]} = e^{-m[1-e^{-a(1-e^{\theta})}]} \quad , \tag{2.30}$$

which is the m.g.f. of the Neyman Type A distribution [Neyman (1939)] with density function:

$$P_r = e^{-m}\frac{a^r}{r!}\sum_{i=0}^{\infty}\frac{i^r}{i!}(me^{-a})^s \quad , \tag{2.31}$$

mean = $m_1 = ma$, and variance = $m_2 = ma(a + 1)$.

The Negative Binomial Distribution. The negative binomial distribution may be derived as a compound Poisson distribution that is compounded by the gamma distribution

$$f_2(x_2) = f_2(\lambda) = \frac{x^k}{\Gamma(k)}e^{-x\lambda}\lambda^{k-1} \quad . \tag{2.32}$$

Thus

$$M_{\lambda a}(\theta) = \int_{-\infty}^{\infty} e^{\lambda a\theta} f_2(\lambda)d\lambda = \int_{-\infty}^{\infty} e^{\lambda a\theta}\,\frac{x^k\lambda^{k-1}e^{-x}}{\Gamma(k)}\,d\lambda \quad ,$$

and, if $y = \lambda x$, then $dy = xd\lambda$, and

$$M_{\lambda a}(\theta) = \int_0^{\infty}\frac{e^{a\theta(y/x)}}{\Gamma(k)}y^{k-1}e^{-y}\,dy$$

$$= \int_0^{\infty}\frac{y^{k-1}}{\Gamma(k)}e^{-y(1-\frac{a\theta}{x})}\,dy$$

$$= \frac{1}{\left(1-\frac{a\theta}{x}\right)^k}\int_0^{\infty}\frac{\left(1-\frac{a\theta}{x}\right)^k}{\Gamma(k)}y^{k-1}e^{-y(1-\frac{a\theta}{x})}\,dy$$

$$= \left(\frac{1}{\underbrace{x - a\theta}_{x}} \right)^k$$

$$= \left(\frac{x}{x - a\theta} \right)^k \qquad (2.33)$$

Whence

$$M_r(\theta) = M_{\lambda a}(e^{\theta} - 1) = \left[\frac{x}{x - a(e^{\theta} - 1)} \right]^k$$

$$= \left(\frac{x}{x + a - ae^{\theta}} \right)^k \quad , \qquad (2.34)$$

which is the m.g.f. of the negative binomial distribution with density function:

$$P_r = \binom{k + r - 1}{r} Q^r P^k \quad , \qquad (2.35)$$

where

$$P = \left(\frac{x}{x + a} \right) \quad \text{and} \quad Q = (1 - P) = \left(\frac{a}{x + a} \right) .$$

2.2.3 Generalized Poisson Distributions

Let X_1 be a Poisson distributed random variable with mean m. Then the random variable $R \equiv X_1 \wedge X_2$ with density function (2.22) is called a generalized Poisson random variable, and its distribution is referred to as a generalized Poisson distribution. Whence

$$F_1(s) = \sum_{x_1=0}^{\infty} e^{-m} \frac{m^{x_1}}{x_1!} s^{x_1} = e^{-m(1-s)} \quad , \qquad (2.36)$$

and

$$F_3(s) = F_1[F_2(s)] = e^{-m[1-F_2(s)]} \quad . \tag{2.37}$$

The Neyman Type A Distribution. The Neyman Type A distribution can also arise as a generalized Poisson distribution in which the generalizer is also a Poisson distribution. Thus, if

$$f_2(x_2) = e^{-a}\frac{a^{x_2}}{x_2!} \quad ,$$

then, by (2.37),

$$F_3(s) = e^{-m[1-e^{-a(1-s)}]} \tag{2.38}$$

and we have, once again, the Neyman Type A distribution of (2.31).

The Negative Binomial Distribution. The negative binomial distribution also can be derived as a generalized Poisson distribution in which the generalizer is the logarithmic distribution. Thus, if

$$f_2(x_2) = \frac{bQ^{x_2}}{x_2} \quad (x_2 = 1, 2, \ldots) \tag{2.39}$$

$$F_2(s) = \sum_{x_2=1}^{\infty} \frac{bQ^{x_2}s^{x_2}}{x_2} = \frac{\log(1 - Qs)}{\log(1 - Q)} = -b \log(1 - Qs) \tag{2.40}$$

$$\text{where } b = -\frac{1}{\log(1 - Q)} \quad ,$$

and, by (2.37),

$$F_3(s) = e^{-m[1+b \log(1-Qs)]}$$

$$= e^{-[mb \log(1-Qs)-m]}$$

$$= e^{[\log (1-Qs)^{-mb} - m]}$$

$$= e^{\log (1-Qs)^{-mb}} \cdot e^{-m}$$

$$= (1 - Qs)^{-mb}(1 - Q)^{mb}$$

$$= \left(\frac{1 - Q}{1 - Qs}\right)^{mb} = \left(\frac{P}{1 - Qs}\right)^{k} \qquad (2.41)$$

where $P = 1 - Q$ and $k = mb$. According to (2.19), this is the p.g.f. of the negative binomial distribution in (2.35).

2.2.4 Other Compound and Generalized Distributions

The above procedures for generating the Neyman Type A and the negative binomial distributions as compound and generalized Poisson distributions can be carried out with many different combinations of distributions to create other new discrete probability distributions. We may, for example, assume other distributions for the compounding or generalizing variable and, by substituting the m.g.f. or p.g.f. of this distribution into (2.29) or (2.37), respectively, generate other compound or generalized Poisson distributions. Thus if the compounder is the rectangular distribution over (α, β), $\alpha < \beta$ then we have the Bhattacharya-Holla (1965) distribution. And if the generalizer is the negative binomial, we obtain the Poisson-negative binomial distribution [Skellam (1952), Katti and Gurland (1961)].

Alternatively, we may wish to deal with compound or generalized distributions that are not Poisson based. Thus, for example, we can compound the n in a binomial distribution with another binomial distribution to obtain yet another binomial distribution. Or, we may generalize a binomial distribution with a negative binomial distribution to derive the binomial-negative binomial distribution [Khatri and Patel (1961)].

PART III

STATISTICAL INFERENCE IN QUADRAT ANALYSIS

3.1 Introduction

Quadrat analysis has both a deductive and an inductive theory. The former strives to determine what frequency distribution tends to be generated by a particular theoretical stochastic process; the latter seeks to identify the stochastic process that is most likely to have generated an observed frequency distribution. We have considered the deductive theory in Part II. Now we shall focus on the inductive theory and examine the related problems of parameter estimation and hypothesis testing.

3.2 Parameter Estimation

To carry out statistical tests on the quality of the fit provided by a theoretical model, to an empirical distribution, it is first necessary to obtain estimates of the theoretical model's parameters. This estimation can be carried out in several ways. The most common procedure is to utilize the method of moments or the computationally more complex method of maximum likelihood.[8]

To estimate the parameters of a distribution by the method of moments one simply equates the sample moments to their theoretical counterparts or, more commonly, one uses the unbiased estimators of the sample moments.

When such moments exist in the theoretical distribution, this method provides consistent extimators of the theoretical distribution's parameters.

---

[8] Other methods such as the method of sample zero frequency and the minimum chi-squared method [Katti and Gurland (1962b)] have been used, and for certain regions in the parameter space these frequently are more efficient than moment estimators. However, they never are more efficient than maximum likelihood estimators.

However, several factors may be cited against the use of moment estimators. First, moment estimators occasionally do not exist. Second, moment estimators are rarely efficient estimators and may lead to erroneous conclusions regarding parametric interrelationships. And, finally, the use of chi-square goodness of fit test is justified only if efficient estimators are used [Chernoff and Lehmann (1954)].

Although the advantages of maximum likelihood estimators are well known, their use in quadrat analysis is rare. Maximum likelihood estimators always exist, although they are sometimes difficult to obtain because the numerical algorithms used to find them occasionally do not converge. Moreover, maximum likelihood estimators are efficient estimators if such estimators exist, and they always are asymptotically efficient.

3.2.1 Moment Estimators

To derive the moments of a theoretical distribution, we recall that given the distribution's probability generating function (p.g.f.):

$$F(s) = \sum_{r=0}^{\infty} P_r s^r \quad ,$$

the following relations hold:

$$F'(1) = \frac{d}{dt} F(1) = \sum_{r=0}^{\infty} rP_r = m_1$$

$$F''(1) = \frac{d^2}{dt^2} F(1) = \sum_{r=0}^{\infty} r(r-1)P_r = m_2' - m_1 = m_2 + m_1^2 - m_1$$

$$F'''(1) = \frac{d^3}{dt^3} F(1) = \sum_{r=0}^{\infty} r(r-1)(r-2)P_r = m_3' - 3m_2' + 2m_1$$

where m_1 is the mean of the theoretical distribution, and m_i' and m_i are, respectively, the i^{th} moment about zero and the i^{th} moment about the mean.

Consequently,

$$m_1 = F'(1)$$

$$m_2 = F''(1) - [F'(1)]^2 + F'(1) \tag{3.1}$$

$$m_3 = F'''(1) - 3F'(1)F''(1) + 2[F'(1)]^3 + 3F''(1) - 3[F'(1)]^2 + F'(1) \quad .$$

If, for example, k_1, k_2, and k_3 are the unknown parameters of the theoretical distribution, they are solutions of the following equations:

$$m_1(k_1, k_2, k_3) = u_1$$

$$m_2(k_1, k_2, k_3) = u_2 \cdot \frac{n}{n - 1}$$

$$m_3(k_1, k_2, k_3) = u_3 \quad \frac{n}{n - 3 + 2/n}$$

where u_1 is the sample mean:[9]

$$u_1 = \frac{1}{n} \sum_r r f_r \ , \quad \left(n = \sum_r f_r \right)$$

u_2 is the sample variance:

$$u_2 = u_2' - u_1^2 = \frac{1}{n} \sum_r r^2 f_r - u_1^2 \quad .$$

and u_3 is the third sample moment about the mean:

---

[9] For convenience, we shall henceforth adopt the notation $\sum_r$ to denote $\sum_{r=0}^{R}$, where R is the value of the largest observed frequency class.

$$u_3 = u_3' - 3u_1u_2' + 2u_1^3 = u_3' - 3u_1u_2 - u_1^3 = \frac{1}{n}\sum r^3 f_r - 3u_1u_2 - u_1^3 \quad .$$

3.2.2 Maximum Likelihood Estimators

The maximum likelihood estimators are those which maximize the likelihood function L.

$$L(k_1, k_2, \ldots, k_p) = \prod_{r=0}^{R} (P_r)^{f_r} \quad ,$$

where R is the largest frequency class observed (i.e., $f_R \neq 0$ and $f_r = 0$ for all $r > R$), P_r is the probability that the result r is observed, and f_r is the number of times the result r was observed in the sample.

To derive the maximum likelihood estimators, we solve the set of equations:

$$\begin{aligned} \frac{\partial}{\partial k_1} \log L &= 0 \\ \frac{\partial}{\partial k_2} \log L &= 0 \\ &\vdots \\ \frac{\partial}{\partial k_v} \log L &= 0 \end{aligned} \qquad (3.2)$$

Equations (3.2) can be written in the more convenient form:

$$\frac{\partial}{\partial k_i}\left[\sum_r \log (P_r)^{f_r}\right] = 0 \qquad (i = 1, \ldots, v)$$

or

$$\sum_r \frac{f_r}{P_r} \frac{\partial}{\partial k_i} P_r = 0 \qquad (i = 1, \ldots, v) \qquad (3.3)$$

In most cases, the roots of (3.3) are difficult to derive, and an iterative procedure is therefore employed. Moment estimators are frequently used as the initial approximations in such methods.

3.2.3 Parameter Estimators for the Component Distributions

The moment estimators for the three component distributions are:

$$\underline{\text{Poisson}}: \quad \hat{m}_1 = u_1$$

$$\underline{\text{Binomial}}: \quad \hat{p} = u_1/n$$

$$\underline{\text{Negative Binomial}}: \quad \hat{m}_1 = u_1 \qquad \hat{k} = \frac{u_1^2}{u_2 - u_1}$$

It can easily be demonstrated that the moment and maximum likelihood estimators for the Poisson and the binomial distributions are identical. This is not true of the negative binomial distribution, however. For the negative binomial distribution in (2.14), with parameters m and k,

$$P_r = \frac{k(k+1) \ldots (k+r-1)}{r!} \left(\frac{k}{m+k}\right)^k \left(\frac{m}{m+k}\right)^r$$

and

$$\frac{\partial}{\partial m} P_r = \frac{r}{m} P_r - \frac{r+k}{m+k} P_r = \frac{k}{m+k}\left(\frac{r}{m} - 1\right) P_r \quad .$$

Equation (3.3) for m therefore is:

$$\sum_r \frac{f_r}{P_r} \frac{\partial}{\partial m} P_r = 0$$

or

$$\frac{k}{m+k} \sum_r \left(\frac{r}{m} - 1\right) f_r = 0 \quad ,$$

and

$$m = \frac{\sum_r r f_r}{\sum_r f_r} = u_1 \quad , \tag{3.4}$$

which is the same result as found by the method of moments.

We have, however, a second parameter, k, and therefore must express the second likelihood equation of the form defined by (3.3):

$$\sum_r \frac{f_r}{P_r} \frac{\partial}{\partial k} P_r = \sum_r f_r \sum_{i=0} \frac{1}{k+i} + \log \frac{k}{k+m} \sum_r f_r + \frac{1}{m+k} \sum_r f_r (m - r) = 0 \tag{3.5}$$

But,

$$\frac{\partial}{\partial k} P_r = \sum_{i=0}^{r-1} \frac{P_r}{k+i} + \log \frac{k}{k+m} + 1 - \frac{k+r}{m+k}$$

and by (3.4) the last term of (3.5) cancels. Rewriting the first term of (3.5) as

$$\sum_r f_r \sum_{i=0}^{r-1} \frac{1}{k+i} = f_1\left(\frac{1}{k}\right) + f_2\left(\frac{1}{k} + \frac{1}{k+1}\right) + \cdots + f_R\left(\frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{k+R-1}\right)$$

$$= \frac{1}{k}(f_1 + f_2 + \cdots + f_R) + \frac{1}{k+1}(f_1 + f_3 + \cdots + f_R) + \cdots + \frac{1}{k+R-1} f_R \quad ,$$

and denoting the sum of observations whose frequency count is greater than r by S_r, i.e.,

$$S_r = f_{r+1} + f_{r+2} + \cdots + f_R \quad ,$$

we obtain from (3.5) the expression:

$$Q(k) = \sum_{r=0}^{R-1} \frac{S_r}{k + r} + n \log\left(\frac{k}{k + m}\right) = 0 \quad . \tag{3.6}$$

Equation (3.6) can only be solved by iterative methods such as the Newton-Raphson method which begins with an initial approximation of k, say k_1, and obtains an improved approximation, say k_2, by means of the expression

$$k_2 = k_1 - \frac{Q(k_1)}{Q'(k_1)}$$

where $Q'(k)$ is the derivative of $Q(k)$. Thus we have

$$Q'(k) = n\left(\frac{1}{k} - \frac{1}{k + m}\right) - \sum_{r=0}^{R-1} \frac{S_r}{(k + r)^2}$$

where a convenient first approximation of k is provided by the moment estimator

$$\hat{k} = \frac{u_1^2}{u_2 - u_1} \quad .$$

To illustrate the differences between results obtained using the two methods of estimation, we present in Table 3.1 both the moment and the maximum likelihood parameter estimates of m and k and the associated expected frequency distributions for the simulated negative binomial point pattern in Figure 2.1b.

TABLE 3.1--OBSERVED AND EXPECTED DISTRIBUTIONS OF THE SIMULATED CLUSTERED (NEGATIVE BINOMIAL) DISTRIBUTION: MOMENT AND MAXIMUM LIKELIHOOD ESTIMATORS

No. of Points per Cell	No. of Cells Observed	N.B. (moment) m = .5200 k = .6781	N.B. (max. like.) m = .5200 k = .7352
0	67	67.98	67.48
1	23	20.01	20.55
2	5	7.29	7.39
3	2	2.82	2.79
4	2	1.13	1.08
5+	1	.78	.70
Total no. of cells = 100			
Total no. of points = 52			
X^2 =		2.05	2.12
u_1 = .5200; u_2 = .9188			
$P_{.05}$ =		5.99	5.99
u_2/u_1 = 1.7669			

3.3 Hypothesis Testing

Let $x_1, x_2, \ldots, x_n$ denote independent observations on a random variable which has an unknown distribution function $F(x)$. The problem of testing whether $F(x) = F_0(x)$, where $F_0(x)$ is some particular distribution function, is called a goodness of fit problem, and any test of the null hypothesis:

$$H_0: F(x) = F_0(x) \quad , \tag{3.7}$$

is called a goodness of fit test.

3.3.1 The Chi-Square Goodness of Fit Test

The most commonly used method for evaluating the fit of a theoretical distribution to an observed one is the chi-square goodness of fit test (hereafter referred to simply as the chi-square test). To carry out this test one groups the empirical data into mutually exclusive discrete frequency classes and then compares this distribution with the expected frequency distribution that would arise from a particular theoretical model. This comparison leads to the calculation of a test statistic which approximately follows a chi-square distribution, only if the hypothesized model is correct. If the model is an incorrect one, the test statistic will tend to exceed the chi-square variate. Thus an ordinary table of chi-square percentiles may be used to determine whether the model provides a satisfactory fit to the data.

The principal advantage of the chi-square test is its simplicity and versatility. It can easily be applied to test the fit of any probability distribution, and it does not require prior knowledge of the distribution's parameters. Its major drawback is its lack of sensitivity, and grouping the data can influence the outcome of the test.

Alternatives to the chi-square test, such as the likelihood ratio test, have not provided it with serious competition for a number of reasons. The most significant reason is that they do not offer advantages not already possessed by the chi-square test. Moreover, a few of them, such as the Kolmogorov-Smirnov test, require a complete specification of the frequency distribution that is followed by the observations. Thus despite the many problems associated with its use, the chi-square test is still the most useful test for goodness of fit that is currently available for quadrat analysis.

3.3.2 The Power of the Chi-Square Test

The Neyman-Pearson theory of hypothesis testing, measures the efficiency of a statistical test by its ability to detect departures from the null hypothesis. Thus, in addition to knowing the random sampling distribution of a given test statistic when the null hypothesis is true, it is also necessary to know the distribution of the same statistic when the null hypothesis is false, and some particular alternative hypothesis is true. Using this latter distribution, one may determine the power of the test.

Assume that $x_1, x_2, \ldots, x_n$ are n independent observations drawn from a population which has an unknown distribution function $F(x)$. Suppose we wish to use the chi-square test to test the null hypothesis

$$H_0: F(x) = F_0(x) \quad .$$

In order to derive the power of this test, it is necessary to specify an alternative hypothesis, say

$$H_1: F(x) = F_1(x) \quad .$$

If H_0 is true, the test statistic χ^2 will exceed χ^2_α, the α significance

point of the chi-square distribution, in a proportion α of the cases. If, however, H_0 is false and H_1 is true, the test statistic will approximately follow the noncentral chi-square distribution, with noncentrality parameter λ.[10]

The power of the chi-square test is the probability that χ^2 exceeds χ^2_α under H_1. Denoting this probability by $p_{df}(\chi^{*2}|\lambda)$, the power function is given by

$$\int_{\chi^2_\alpha}^{\infty} p_{df}(\chi^{*2}|\lambda)d\chi^{*2} \equiv \beta(df, \lambda, \alpha) .$$

Thus the power for a given number of degrees of freedom, df say, and significance level, α is a monotonically increasing function of the noncentrality parameter λ. Therefore, we may express the null hypothesis as

$$H_0: \lambda = 0$$

against the alternative hypothesis

$$H_1: \lambda > 0 ,$$

where H_1 is a composite hypothesis.

Let us illustrate the computation of the power function with an example drawn from quadrat analysis. In such analyses it is frequently the case that spatial distributions that are well fitted by the Neyman Type A distribution are also satisfactorily accounted for by the negative binomial distribution. Thus we frequently end up dealing with the null hypothesis that the empirical distribution is a Neyman Type A distribution against

---

[10] Tables of the noncentral chi-square distribution have been published by Fix (1949) and Patnaik (1949).

the alternative hypothesis that it is a negative binomial distribution. We have then that

$$H_0: F(x) = \text{N.T.A. } (m, a)$$

and (3.8)

$$H_1: F(x) = \text{N.B. } (m, k) \quad .$$

We have seen that the power function depends on three parameters df, λ, and α. Therefore, for a given sample size, say n = 400, and a pre-specified significance level, say $\alpha = .05$, we can express power as a function of λ. However, in light of our hypothesis in (3.8) and in recognition of our concern with spatial analysis, a more useful formulation is one which expresses power in terms of the mean and the variance-mean ratio of a distribution. Such a measure is particularly relevant in that it provides an indication of the degree of clustering which is present in the empirical distribution being analyzed. In order to replace λ with the mean and the variance-mean ratio, it is necessary to compute, for each mean and variance-mean pair, a value of λ and then, using that value, to associate with this particular pair, the approximate power of the chi-square test to discriminate between the two hypothesis expressed in (3.8). Thus, we proceed as follows:

(1) Given the population mean $m_1 = m_1^*$ and variance $m_2 = m_2^*$, compute the variance-mean ratio m_2^*/m_1^*.

(2) Compute the expected frequency distribution of the Neyman Type A distribution for $m_1 = m_1^*$, $m_2 = m_2^*$, and n = 400.

(3) Compute the expected frequency distribution of the negative binomial distribution for the same values of m_1, m_2, and n.

(4) Compute the noncentrality parameter $\lambda = \lambda^*$ and determine the degrees of freedom by subtracting 1 from the number of frequency classes R.

(5) Using a table of the noncentral chi-square distribution such as Patnaik's, find the power β^* which is associated with $\alpha = .05$, $\lambda = \lambda^*$, and df = R - 1.

Carrying out these calculations with a numerical example in which m_1, in sequence, takes on the values of .111, .444, 1.000, 1.778, 2.778, and 4.000, respectively, while at the same time m_2/m_1, in the same sequence, takes on the values of 1.5, 2.0, 2.5, 3.0, 5.0, and 10.0, we may construct Figure 3.1.

Figure 3.1 illustrates the following very important property of the chi-square test: For a given value of the mean, the power of the test to discriminate between a Neyman Type A distribution and a negative binomial distribution increases with the variance-mean ratio. Since the negative binomial is, in a sense, a more "clustered" distribution than the Neyman Type A, it is intuitively apparent that the more clustered the empirical point pattern is, the easier will it be to reject the null hypothesis, in (3.8), in favor of the alternative hypothesis. A corollary of this finding is that for (3.8) increasing the variance, while holding the mean constant, increases power.

3.3.3 Some Problems in the Use of the Chi-Square Test

Any competent application of the chi-square test must recognize the conditions which restrict the usefulness of the test. One of the more important of these has already been alluded to in our discussion of power. There we noted that chi-square tests commonly are not directed against a specific alternative hypothesis. Hence they cannot point to the way in which the data and the null hypothesis differ, although such differences may be particularly informative. Further, we have seen that the power of the chi-square test to detect a mismatch between theory and data depends on the size of the sample. With a small sample, it is very difficult to discriminate

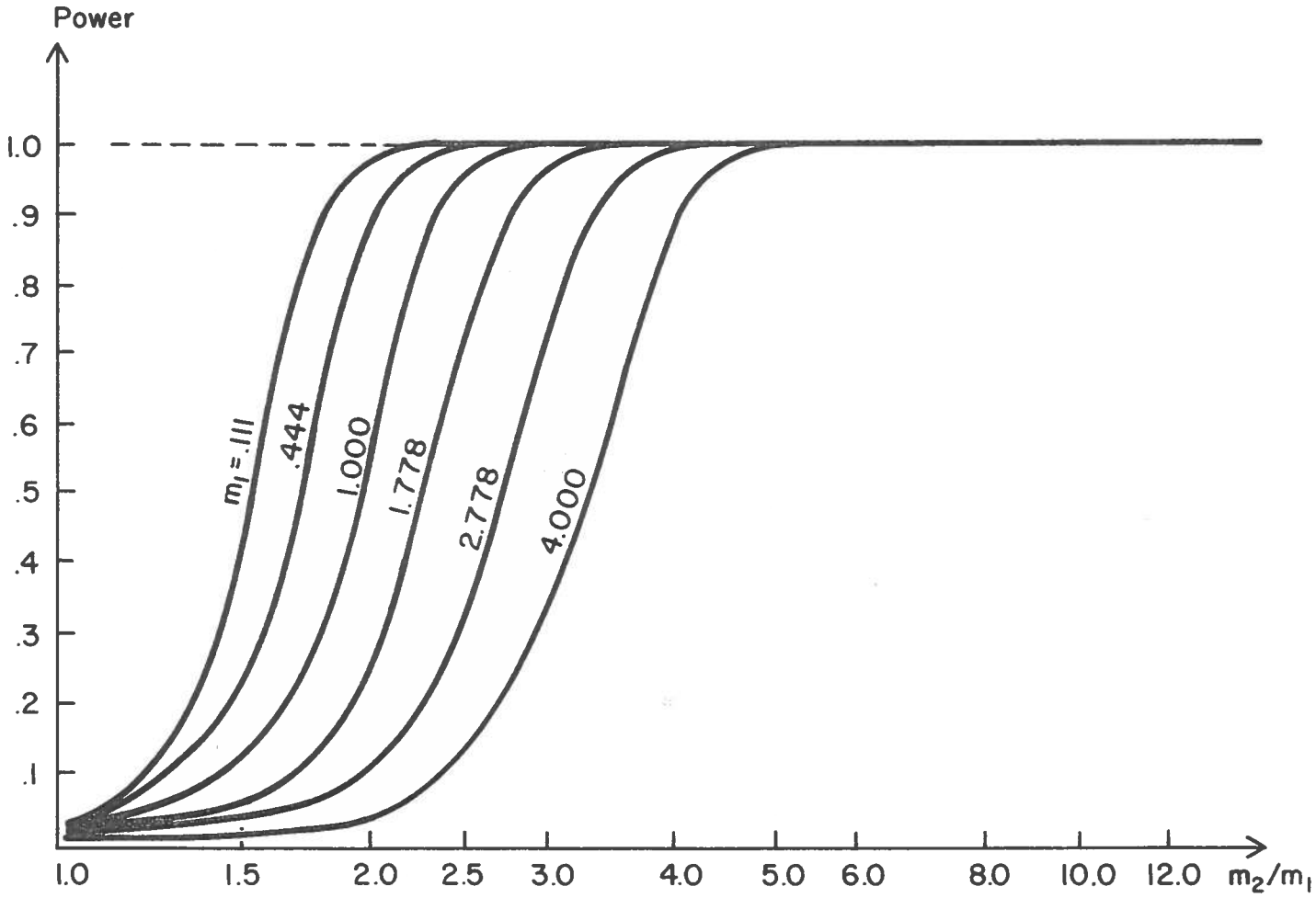


Fig. 3.1--APPROXIMATE POWER FUNCTION OF THE CHI-SQUARE TEST FOR n=400,

α=.05 : N.T.A. vs. N.B.

between a null hypothesis and an alternative one, because the latter is unlikely to produce a significant value of X^2 in such cases. In a very large sample, however, small departures from the null hypothesis are almost certain to be identified. Thus, in instances of a failure to reject the null hypothesis, our confidence in this hypothesis is directly a function of sample size.

Two other important problems warrant attention. These grow out of the estimation procedure that is used to obtain parameter estimates and the need to group small expectations prior to the computation of the test statistic.

Estimation Problems. Fisher (1924) pointed out that the limiting distribution of the X^2 statistic depends on the method of estimation used to obtain it. An inefficient estimator may lead to a large value of X^2, even if the theory is correct. Fisher also showed that, for large samples, the method which chooses the unknown parameters so as to minimize the value of X^2, in the limit, becomes equivalent to the method of maximum likelihood. His argument leads to two important results. Any efficient estimator yields estimates which asymptotically are identical to the maximum likelihood estimates, and the χ^2 distribution, with the appropriate reduction in the degrees of freedom, is valid only when efficient estimators are used.

Chernoff and Lehmann (1954) extend Fisher's results. They point out that, when the values of the individual observations are available, it is more efficient to utilize this knowledge in the parameter estimation process, even though one ends up aggregating the results into the R class frequencies before calculating the test statistic. They prove, however, that when the X^2 statistic is obtained with maximum likelihood estimation that is based on the n individual observations, instead of the R frequencies,

f_r, it no longer has an asymptotic chi-square distribution. Its true distribution is bounded between a χ^2_{v-1} and a χ^2_{v-s-1} variable, and as v increases, the difference decreases sufficiently rapidly in order for this difference to be ignored. For small v, however, the use of the χ^2_{v-s-1} distribution for test purposes may lead to serious error in that the probability of exceeding any given value will be greater than we suppose. Kendall and Stuart (1961) suggest, therefore, that in such situations the test statistic should exceed the critical values of both distributions.

Grouping Problems. In order to prove that X^2 is asymptotically distributed as χ^2 under the null hypothesis, it is necessary to assume that the expectations, nP_{0r}, tend to infinity. For this reason, it is customary, in practical applications of the chi-square test, to recommend that the expected number of observations in any frequency class should not be too small. Combination of neighboring classes to satisfy this constraint, therefore, is generally carried out. However, there is little agreement about the admissible minimum value that should be used. Kendall and Stuart (1961), for example, suggest that all nP_{0r} should be not less than 10. Another lower bound that seems to be popular with statisticians is 5. Cochran (1954), however, argues that in certain situations it is possible to reduce this minimum to 1, without serious loss of accuracy.

Several difficulties arise as a consequence of grouping. First, the pooling of observations may result in there not being enough frequency classes for a chi-square test to be carried out. That is, there may not be any degrees of freedom left after the reduction for parameter estimates is taken into account. Second, grouping may hide significant deviations between observed and the oretical distributions. For example, grouping may obscure discrepancies from a Poisson distribution and thereby fail to identify

nonrandomness when it exists. This may be illustrated by considering the following numerical example:

Number of points per quadrat	Observed number of quadrats	Expected number of quadrats under H_0: F(x) = Poisson (m_1 = .52)
0	67	59.45
1	23	30.92
2	5	8.04
3	2	1.39
4	2	.18
5+	1	.02

The above observed frequency distribution is a clustered one, as is evidenced by its variance-mean ratio which is greater than 1, and its X^2 statistic, with two degrees of freedom, which is significant at the 5 per cent level (X^2 = 11.40). Pooling the last 4 frequency classes to maintain a minimum expectation bound of 5, we obtain:

Number of points per quadrat	Observed number of quadrats	Expected number of quadrats under H_0: F(x) = Poisson (m_1 = .52)
0	67	59.45
1	23	30.92
2+	10	9.63

As a consequence of the above grouping, the X^2 statistic now is equal to 3.00, with 1 degree of freedom. Thus we fail to reject the hypothesis of randomness at the 5 per cent level of significance.

We may generalize the above observations by concluding that grouping reduces power. Loss of power occurs because grouping generally is carried out on classes at the tails or extremes of a distribution. These frequently are the very places where the data deviate most clearly from the hypothetical frequencies. Information about the extent of the loss of power is difficult

to obtain, because the power function of the chi-square test is known only as a limiting result, a result which assumes that the expectations are large. Nevertheless, an approximate measure of power may be obtained by using tables of the noncentral chi-square distribution.

3.3.4 The Problem of Quadrat Size

Ecologists have pointed out that the size of the square which is used to sample a study area has a considerable impact on the findings that are generated by quadrat analysis [Greig-Smith (1962)]. The appearance of non-randomness in a sample of quadrats is not an absolute characteristic, but depends on the size and shape of the sample unit used. Thus, for example, a clustered pattern sampled with a relatively small quadrat generally will yield findings that are quite different from those that would be indicated by a sample drawn using a relatively large quadrat.

The shortcomings of the quadrat method have become increasingly apparent in recent years as the ways in which quadrat size affects spatial data have been explored much more precisely. Much of the ecological literature on this subject has been reviewed by Curtis and McIntosh (1950), who suggest the general rule of thumb that the appropriate size of square is approximately twice the size of the mean area per point. Greig-Smith (1964), on the other hand, suggests that minimizing the sample variance and striving for a reasonably symmetric distribution are appropriate criteria for certain classes of distributions.

Rogers and Gomar (1969) suggest an approach for determining the optimal quadrat size which is more in keeping with the Neyman-Pearson theory of statistical inference. They propose that the optimal quadrat size is that size which maximizes the power of the chi-square test for a given level of significance, α. This criterion, of course, requires that one specify not only the null hypothesis, but the alternative hypothesis as well.

PART IV

QUADRAT ANALYSIS OF URBAN RETAIL SPATIAL STRUCTURE

4.1 The Spatial Structure of Retailing in Urban Areas

Retail activity is concerned with the distribution of goods and services to their ultimate consumers--households. Retail establishments are business firms which perform this function for the population in the area surrounding them, i.e., their market area. This territorial delineation stems from the interaction of two components: the kinds of consumer goods and services that are being marketed, and the spatial behavior of consumers purchasing these goods and services. Their interaction defines the areal domain from which establishments are able to attract a clientele. A few theorists [Berry and Garrison (1958a,b)] have labeled the upper limit of the radius of this domain as the "range" of a good and have combined this notion with the concept of a minimum "threshold" sales level to account for the hierarchical arrangement of retail and service centers in urban areas.

4.1.1 Classes of Retail Goods

Marketing specialists typically distinguish between three classes of consumer goods: convenience goods, shopping goods, and specialty goods. The standard definitions of these terms reflect the buying habits of the consumer and typically distinguish goods on the basis of the frequency of their purchase and the role brand names play in the purchasing process.

Convenience goods are consumer goods which are brought frequently and repeatedly, and which the consumer therefore desires to purchase with a minimum of effort. Brand names, rather than price, are the principal consideration. Examples are groceries, tobacco, and drug items. Since these

goods are purchased day after day, convenience and accessibility are of fundamental importance. Thus, stores carrying these goods tend to locate close to the consumer.

Shopping goods are generally acquired at periodic intervals: seasonally, annually, and, in some instances, only once in a lifetime. Such items commonly are unstandardized and are purchased only after considerable comparisons of quality and price have been made in a number of competing retail outlets. Furniture, shoes, jewelry, and style goods in general are prominent examples of this category of consumer goods. To satisfy the consumer's desire to "shop around" shopping goods stores tend to cluster in major retail centers.

Specialty goods are goods for whose purchase the consumer is willing to make a special purchasing effort. Price once again is not the major consideration, but plays a subordinate role to merchandise quality. Examples are expensive perfumes, watches of superior quality, and high-grade men's and women's clothing. Since consumers of specialty goods typically make a special visit to purchase them, the locational decisions of stores handling such merchandise are quite free of any restraining conditions such as are met by convenience goods or shopping goods outlets. Thus specialty goods stores may locate virtually anywhere in the urban area without seriously affecting their volume of sales.

It is quite apparent that a considerable degree of overlap exists in the above classification system. Different consumers may as a result of their buying habits treat convenience goods as specialty goods and vice versa. Thus, though most people consider groceries as a convenience good, many "shop around" for them. On the other hand, common work shoes may be treated as a convenience good and purchased at the most convenient location.

These difficulties of overlapping categories are inherent in any classification system and once recognized should not seriously hamper most analytical investigations.

4.1.2 Intraurban Retail Spatial Structure

Despite the great physical changes brought about by the tremendous growth of suburban areas, the retail spatial structure of most urban areas has maintained a remarkably constant profile during the past thirty years. Retail establishments in most major cities generally belong to one of the following classes of centers: the central business district, outlying secondary shopping districts, neighborhood business nucleations or small retail clusters, and isolated individual stores.

Retail spatial patterns arise directly out of the locational decisions of the distributors of retail goods and services. The individual business firm operating with incomplete information regarding the urban environment seeks to serve the spatially scattered market by establishing itself at a profitable location. Analyses of the significant influences bearing on its locational decision form the body of retail location theory.

Virtually all of the theories that attempt to analyze, interpret, and ultimately predict the spatial behavior of retail establishments point to a common set of factors which together are held to influence the locational decisions made by such firms. These are variables having as their principal referent the consumer, the state of the urban system, or the establishment. Related to the first are factors such as the distribution of population, income, and purchasing power. Connected with the second are considerations pertaining to accessibility, site economics, and the availability of advantageous locations. Finally, falling in the third category are factors which reflect the establishment's accommodation of internal needs to the

external setting, as for example, the affinity for other economic activities and the role of spatial monopoly in competition.

The general rationale that emerges concerning the spatial behavior of retail outlets is that the geographical configuration of retail and service trade establishments in urban areas reflects the averaging of the attracting and repelling forces of linkages incident to the operation of such firms. Accessibility to the residentially based market, exposure to vehicular or pedestrian movements, and proximity to competing and complementary establishments are typical examples of some of the fundamental considerations that appear to be present in these locational decisions.

4.1.3 Competition

The theory of retailing and service trade describes retailing as monopolistic or imperfect competition [Aubert-Krier (1954), Lewis (1945)]. Although many features of the structure of retail markets are included in the generally accepted definitions of perfect competition, the presence of certain "imperfections" has led writers to place retail activities in the middle ground between perfect competition and monopoly. These imperfections stem from two major sources: 1) spatial and nonspatial monopolistic influences such as favorable locations or prestige clientele; and 2) the desire of retail firms to minimize risk and limit the area of competitive activities, such as price competition, that can be undertaken by other firms.

Retail establishments continually evaluate the impact that their decisions will have on competitors and generally pay close attention to the policies of other establishments carrying similar lines of merchandise within the same geographical trading area. This is in contrast to "perfect competition" where the behavior of a firm has no significant influence on market price. Also, it is frequently held that since people tend to shop

around for shopping goods price elasticity of demand is much greater for shopping goods than convenience goods and, as a result, shopping goods retailing is competitive. The same poing of view suggests that since consumers attempt to minimize the disutility of convenience goods shopping, and tend to purchase many convenience goods at one place, convenience goods retailing tends to be imperfectly competitive or monopolistic. This is an oversimplified view of the fundamental difference between the market structure of the two branches of retail trade. However, it does appear to have some merit in explaining the spatial concentration of particular classes of retail firms and the decentralization of others.

4.1.4 The Clustering and Dispersal of Retail Establishments in Urban Areas

The tendency for particular classes of retail establishments to cluster has been noted by several theorists on the subject of intraurban retail spatial structure. Hotelling (1929), for example, in his classical essay on spatial competition, concludes that as "more and more sellers of the same commodity arise, the tendency is not to become distributed in the socially optimum manner but to cluster unduly." Nelson (1958) supplies a rationale for this tendency with his "principle of cumulative attraction." This principle holds that a group of stores carrying the same merchandise will do more business if they are clustered together than if they are widely scattered. Thus, for example, a nucleus of three shoe stores will do a larger volume of business than will the same three located several blocks apart.

Reference to the generally regular pattern of convenience goods outlets in urban areas also appears in much of the theoretical literature on retail spatial structure. Ratcliff (1939), for example, recognizes that certain neighborhood type establishments, such as grocery stores, do not

profit from cohesion and therefore typically do not group together. Others assert that the scattered pattern of convenience goods stems from their limited "range" and the restricted market areas that they command as a result of this characteristic [Berry and Garrison (1958a,b)]. Finally, marketing researchers such as Nelson (1958), note that a convenience goods establishment, in selecting a location, should be primarily concerned with accessibility and apparent freedom from future competition.

It appears then that explanations of the spatial pattern of consumer goods firms typically cite two interdependent factors as the principal contributory agents to clustering or regularity: the nature of consumer goods, and the structure of consumer goods markets. Catering to the immediate every day needs of the consumer, convenience goods establishments have followed the urban population to the outlying areas of metropolitan communities and have assumed a spatial pattern which mirrors that of the population they serve. The day after day needs for such goods have made accessibility and proximity of fundamental importance in the purchasing process, and the hazards of competition encourage dispersal. Shopping goods establishments, however, frequently reflect the thesis that firms with a high degree of compatibility can ensure a greater volume of sales for themselves by clustering together. Their compatibility may stem from the complementary nature of their activities and products, or may occur when, though competitive, they carry goods of different quality and price. Thus, in a very general sense, we may conclude that rival shopping goods establishments tend to attract one another and thus to assume a more regular spatial pattern. Establishments falling outside this dichotomous classification may be viewed as producers of specialized goods and services and may be assumed to have locational characteristics that are appropriate for their particular activity. Typical

examples are traffic oriented activities, such as gasoline service stations, which appear to be highly dependent on the traffic flow pattern of an urban area.

4.2 Compound and Generalized Distributions as Models of Retail Spatial Dispersion

Compound and generalized distributions offer a variety of models that can be fitted to empirical data. Both classes of models assume that a given urban area has been gridded into a network of squares, and both account for the observed dispersion in this network by describing the spatial regularity or clustering of retail establishments as a realization of a general stochastic process. Compound distributions postulate areal inhomogeneity of some important parameter such as the mean. Generalized distributions reflect the assumption that stores occur only in clusters.

4.2.1 The Compound Model

Assume that in an urban area which has been gridded into a network of squares the number of stores in each square, X_1 say, may be described by the probability distribution $f_1(x_1)$. Next, assume that some parameter of $f_1(x_1)$, X_2 say, varies from place to place, within the study area, according to the probability distribution $f_2(x_2)$. Then $f_1(x_1)$ is a conditional probability distribution that depends on the values assumed by X_2 and, therefore, can be denoted as $f_1(x_1|x_2)$. Thus P_r, the unconditional probability distribution of the number of stores in each quadrat, may be expressed as

$$P_r = \int_{-\infty}^{\infty} f_1(r|x_2) f_2(x_2) dx_2 \quad , \tag{4.1}$$

which, except for the missing arbitrary constant c, is the compound distribution defined in (2.20).

4.2.2 The Generalized Model

Given an urban area that has been gridded into a network of quadrats, consider a model that seeks to describe the spatial dispersion of stores within this network by means of a general stochastic process based on the following postulates:

1. Stores occur only in clusters.

2. The number of clusters in a quadrat, X_1 say, is described by the probability distribution $f_1(x_1)$.

3. The number of stores in a cluster, X_2 say, varies from one cluster to another according to the probability distribution $f_2(x_2)$, which is the same for all clusters.

The statement of theproblem of retail spatial dispersion in these terms identifies the process as a generalized distribution with the probability generating function defined in (2.27).

4.3 Quadrat Analysis of the Spatial Dispersion of Retail Establishments in Urban Areas

Two sets of data will be used for the empirical testing of the distributions that have been defined in this paper. In particular, we shall focus on the fits provided by the binomial, the Poisson, the negative binomial and the Neyman Type A distributions to data describing the spatial pattern of retail establishments in Ljubljana, Yugoslavia and San Francisco, California. The former were compiled for the author by Lidja Podbregar-Vasle of the Town Planning Institute in Ljubljana; the latter were made available to the author by the staff of the Bay Area Transportation Study Commission, with the permission of California's State Department of Employment.

4.3.1 Interurban Comparisons

Ljubljana is the capital of the Slovenian Republic in Yugoslavia. The city is built on the site of the Roman town of Emona, which was destroyed early in the 5th century. It has a population of about 200,000 and, along with Zagreb in Croatia, serves as the economic, cultural and political center of northern Yugoslavia.

Figure 4.1 presents the spatial pattern of retail establishments in Ljubljana during the year 1966. Of the 440 retail establishments inside of the square 3,000 by 3,000 meter study area, 201 are food stores and, of these, 79 are grocery stores (the remainder being fruit and vegetable shops, butcher shops, stores selling only bread and milk, confectionaries, etc.). Of the 239 nonfood stores, 68 stores have been classified as clothing, i.e., apparel, stores (stores primarily engaged in selling clothing, shoes, hats, etc.). Figure 4.2 presents the spatial patterns of these four categories of retail establishments.

The study area may be gridded in a number of different ways for purposes of quadrat analysis. The 3,000 by 3,000 meter square can be viewed as a 60 by 60 grid of squares, 50 meters on a side, and then these may be aggregated to form 30 by 30, 20 by 20, 15 by 15, 12 by 12 and 10 by 10 grid systems. For each of these observed series, parameter estimates may be derived by the method of maximum likelihood, and for each distribution the relevant expected values can be calculated. The adequacy of the fit provided by each of the theoretical frequencies then may be tested by means of the χ^2 goodness of fit test. To eliminate the disproportionate effect of small differences from a low expectation, let us pool frequencies with small expectations so that no expectation is less than 5. And to condense an otherwise unwieldy array of results, we shall only consider the fits provided

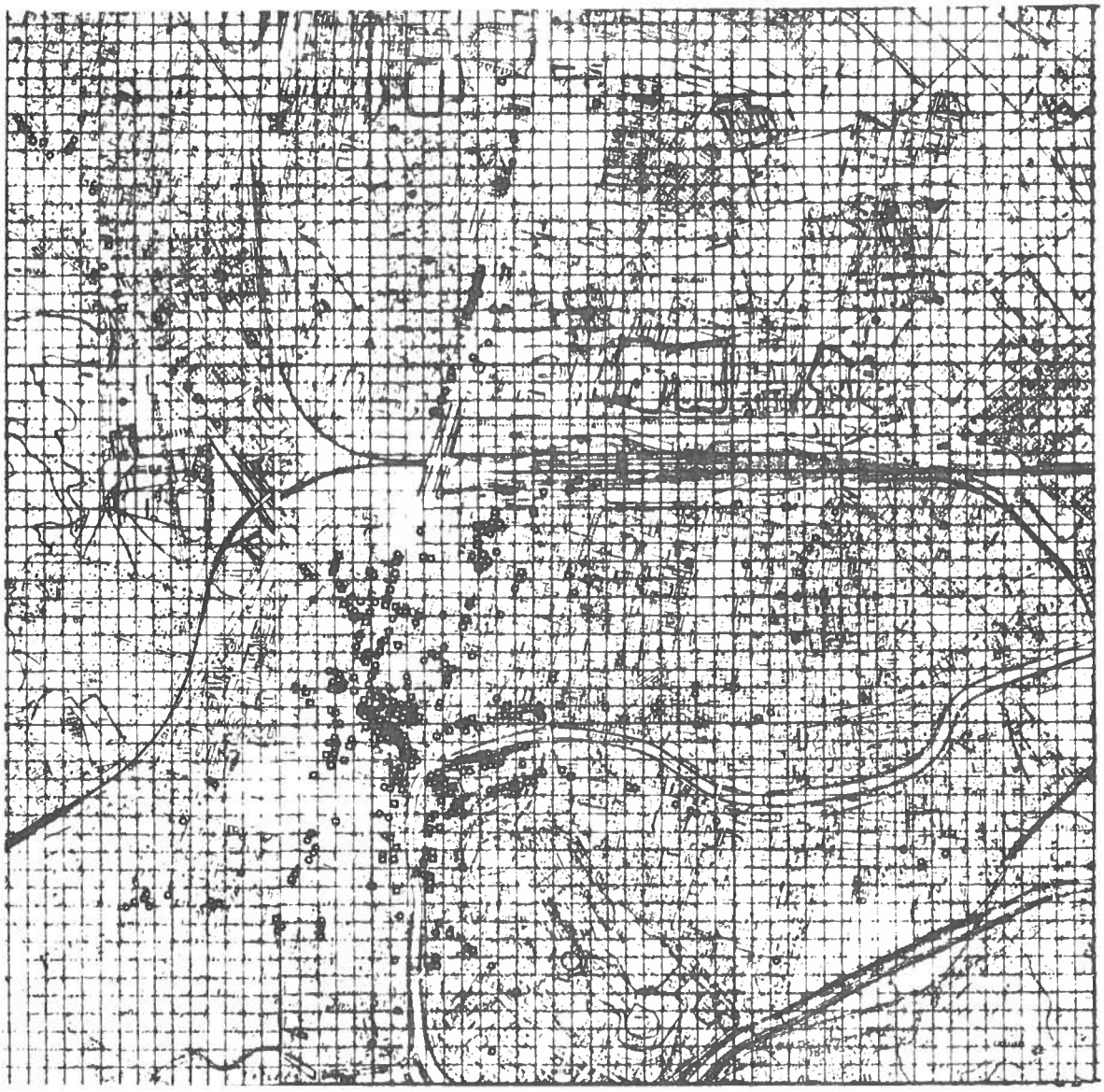


FIG. 4.1--THE SPATIAL PATTERN OF RETAIL ESTABLISHMENTS IN LJUBLJANA, YUGOSLAVIA

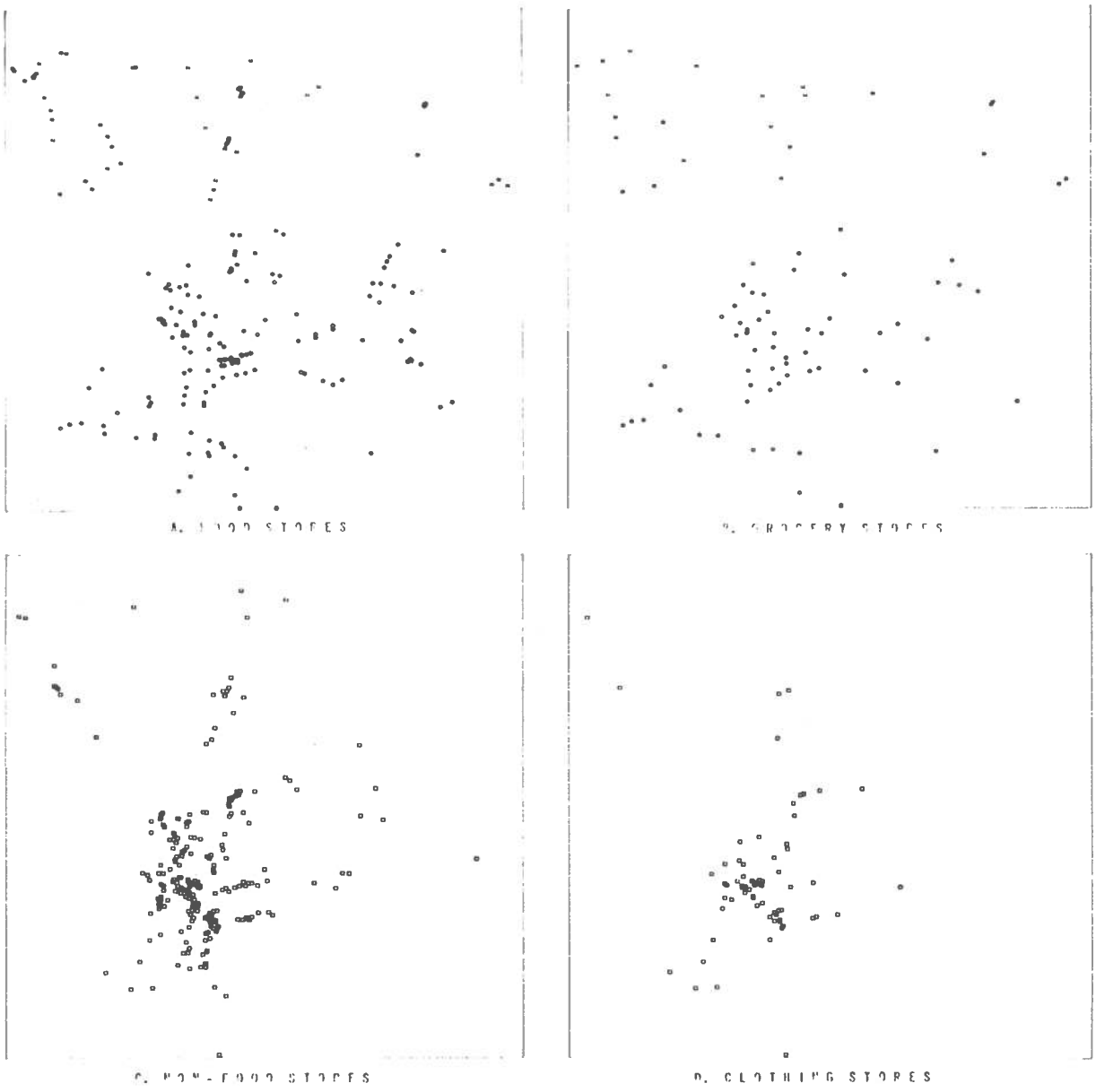


FIG. 4.2--THE SPATIAL PATTERNS OF FOOD, NONFOOD, GROCERY AND CLOTHING STORES IN LJUBLJANA, YUGOSLAVIA

TABLE 4.1--PARAMETER ESTIMATES AND THE GOODNESS OF FIT FOR VARIOUS CLASSES OF RETAIL ESTABLISHMENTS IN LJUBLJANA, YUGOSLAVIA

a. Total stores

Model	$u_1 = 3.0556$ $u_2 = 53.0458$ $u_2/u_1 = 17.3605$	X^2	d.f.
Binomial	n = 51 p = .0559	898.20*	5
Poisson	m = 3.0556	831.15*	6
Negative Binomial	m = 3.0556 k = .2255	4.41	6
Neyman Type A	m = .7765 a = 3.9351	46.27*	7

b. Food stores

Model	$u_1 = 1.3958$ $u_2 = 6.9261$ $u_2/u_1 = 4.9620$	X^2	d.f.
Binomial	n = 22 p = .0634	139.10*	3
Poisson	m = 1.3958	123.09*	3
Negative Binomial	m = 1.3958 k = .3544	3.99	4
Neyman Type A	m = .6435 a = 2.1691	4.14	4

c. Grocery stores

Model	$u_1 = .5486$ $u_2 = 1.0745$ $u_2/u_1 = 1.9587$	X^2	d.f.
Binomial	n = 5 p = .1097	8.64*	1
Poisson	m = .5486	5.47*	1
Negative Binomial	m = .5486 k = .6374	1.49	1
Neyman Type A	m = .8975 a = .6112	2.61	1

d. Nonfood stores

Model	$u_1 = 1.6597$ $u_2 = 29.7086$ $u_2/u_1 = 17.8997$	X^2	d.f.
Binomial	n = 43 p = .0386	304.75*	3
Poisson	m = 1.6597	290.26*	3
Negative Binomial	m = 1.6597 k = .1073	3.49	3
Neyman Type A	m = .4205 a = 3.9475	39.84*	5

e. Clothing (Apparel) stores

Model	$u_1 = .4722$ $u_2 = 3.8174$ $u_2/u_1 = 8.0839$	X^2	d.f.
Binomial	n = 16 p = .0295	29.76*	1
Poisson	m = .4722	28.53*	1
Negative Binomial	m = .4722 k = .1038	3.53	1
Neyman Type A	m = .2426 a = 1.9464	12.31*	2

---

* = X^2 statistic significant at 5 per cent level.

TABLE 4.2--OBSERVED AND EXPECTED DISTRIBUTIONS OF FOOD STORES IN LJUBLJANA, YUGOSLAVIA

No. of stores per cell	No. of cells observed	B.	P.	N.B.	T.A.	P.B.(2)	P.B.(3)	P.B.(4)	P.B.(5)	P.N.B.
0	83	34.05	35.66	81.76	81.44	61.37	71.59	75.11	77.16	82.89
1	18	50.74	49.77	23.11	12.99	19.03	13.47	12.83	12.41	19.85
2	13	36.09	34.74	12.48	15.12	36.26	24.49	20.61	18.78	12.57
3	9	16.30	16.16	7.81	12.49	10.64	17.79	16.59	15.65	8.50
4	7	5.25	5.64	5.22	8.55	10.67	6.69	8.42	8.89	5.89
5	7	1.28	1.57	3.63	5.37	2.97	5.30	4.64	4.83	4.13
6	2	.25	.37	2.58	3.29	2.08	2.55	2.93	2.93	2.92
7	1	.04	.07	1.87	1.99	.55	1.09	1.47	1.66	2.07
8	2		.01	1.37	1.18	.30	.60	.72	.85	1.48
9	0			1.01	.69	.08	.25	.36	.43	1.05
10+	2			3.16	.89	.05	.18	.32	.41	2.63
Total no. of cells = 144	X^2 = 139.10	123.09	3.99	4.14	34.62	14.71	10.66	8.72	1.66	
Total no. of stores = 201	$P_{.05}$ = 7.82	7.82	9.49	9.49	7.82	7.82	7.82	7.82	7.82	

by one quadrat size, the "optimal" size which maximizes power in the sense defined in Section 3.3.4. Using the simulated power functions reported in Rogers and Gomar (1969), we find that the 12 by 12 grid system, with 250m^2 quadrats, most nearly satisfies this requirement for our data. Thus in Table 4.1 we present the fits of the binomial, the Poisson, the negative binomial and the Neyman Type A distributions to the data on the spatial distribution of Ljubljana's total, food, nonfood, grocery and clothing (apparel) stores, when sampled with a square grid of 144 quadrats 250 meters on a side. Table 4.2 provides more detailed information for one class of establishments: food stores. In it we present a complete list of observed and expected frequencies and include the fits of two other generalized Poisson distributions: the Poisson-binomial (n = 2, 3, 4, 5) and the Poisson-negative binomial.

For each of the five empirical distributions of retail establishments investigated, the negative binomial distribution clearly provides the most satisfactory fit. Of the four null hypotheses examined in each instance, it alone is not rejected at the 5 per cent level of significance. Thus, one may with reasonable justification, conclude that this theoretical distribution provides an adequate accounting for the observed data.

The theoretical model generating the negative binomial, however, cannot be clearly identified. As we have seen, there are a number of ways in which the negative binomial distribution can arise. One cannot, therefore, on the basis of an observed set of data alone, distinguish between the different probability mechanisms that may be operating to produce this distribution. Thus, for example, it has been shown, in Section 2.2.3, that if clusters of establishments are randomly distributed and the number of establishments in each cluster follows a logarithmic distribution, then the resulting number

of stores per cell is given by the negative binomial distribution. It also has been demonstrated, in Section 2.2.2, that the same distribution will arise if the number of stores per cell does follow a random distribution, but one in which the mean is allowed to vary according to a gamma probability distribution.

It is highly probable that both mechanisms are operative in the retail activity system. Retailers carrying the same class of shopping goods merchandise do indeed appear to be attracted to one another; however, they also are attracted to purchasing power. Since purchasing power is unevenly distributed across the urban landscape, one may with some justification assert that the number of retail establishments that locate in a particular cell is distributed according to a Poisson distribution, but not necessarily with the same mean. That is, it could be held that the average number of stores per cell varies according to the purchasing power present in that cell. Thus, if purchasing power is distributed according to a gamma distribution, the negative binomial distribution describes the number of stores per cell.

Recall that the negative binomial distribution is one of a series of distributions that may be used to account for data in which the variance is larger than the mean. The variance of the negative binomial is

$$m_2 = m_1 + \frac{m_1^2}{k} \quad ,$$

where m_1 is the mean and k the exponent. Hence, it is clear that as $k \to \infty$, the variance m_2 approaches the mean m_1. That is, as k increases, the ratio of the variance over the mean approaches unity. Indeed, in Section 2.1.4, we have shown that in such instances the distribution itself converges to the Poisson distribution. On the other hand, as $k \to 0$ from the positive direction the variance increases without bound and greatly exceeds the mean indicating a highly clustered distribution.

It appears, therefore, that the relative value of the parameter k in the negative binomial distribution may be used to indicate the "spatial affinity" that firms in a particular class of retail establishments have for one another. For the Ljubljana data, the following estimates for k were derived:

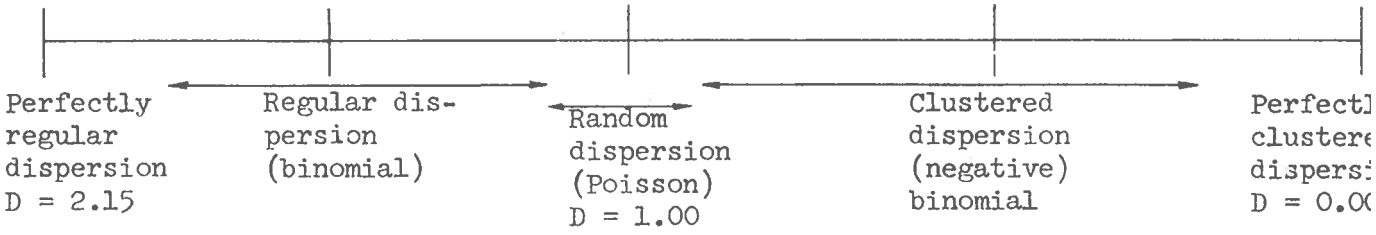
Clothing (Apparel) stores:	k = .1038
Nonfood stores:	k = .1073
Total stores:	k = .2255
Food stores:	k = .3544
Grocery stores	k = .6374

Thus, clothing stores seem to exhibit the highest degree of clustering and grocery stores the least. Almost identical findings are indicated by the corresponding nearest-neighbor statistics, $\bar{d}_0$ and D, defined in (1.11) and (1.14), respectively:

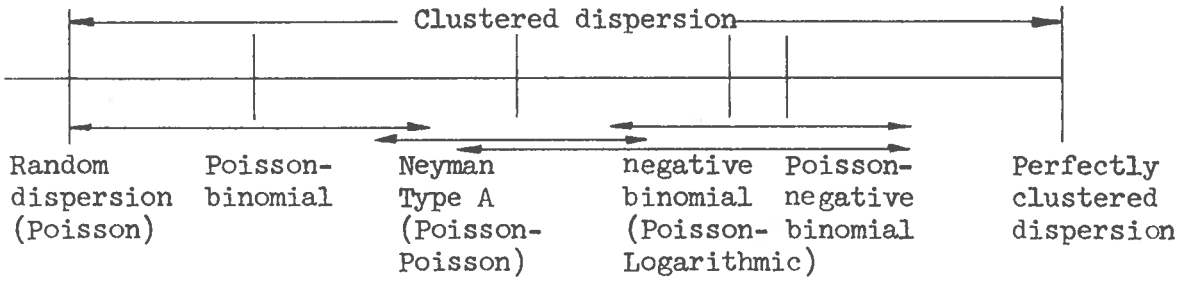
Clothing (Apparel) stores:	$\bar{d}_0 = 0.80$	D = .4393
Nonfood stores:	$\bar{d}_0 = 0.37$	D = .3842
Total stores:	$\bar{d}_0 = 0.32$	D = .4547
Food stores:	$\bar{d}_0 = 0.54$	D = .5091
Grocery stores:	$d_0 = 1.46$	D = .8674

The position of clothing and nonfood stores would be reversed if D were the measure used to indicate "clusterdness."

Notice that the spatial dispersion of food stores and grocery stores in Ljubljana also can be described by the Neyman Type A distribution. How can we interpret this finding? First, we note that the range of distributions between perfectly regular and perfectly clustered dispersions may be illustrated with the following dispersion line:



Next, focusing only on dispersions that are more clustered than random, let us locate the Neyman Type A, the Poisson-binomial and the Poisson-negative binomial distributions on the right side of the dispersion line, according to their degree of "clusterdness":[11]

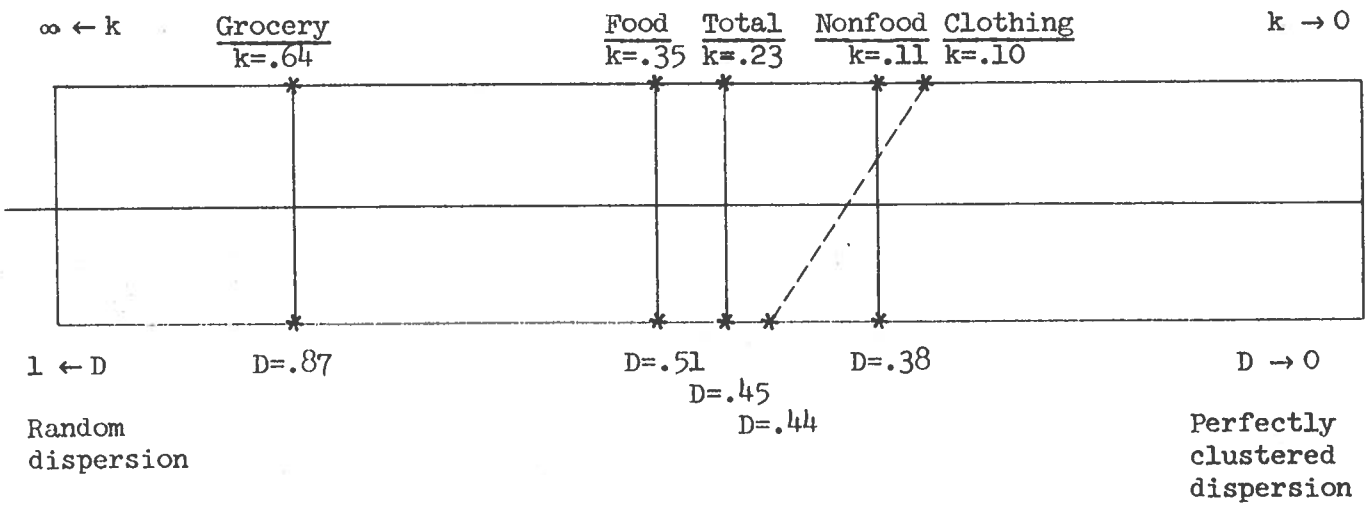


We conclude, therefore, that the reason both the negative binomial and the Neyman Type A distributions can be used to describe the spatial dispersion of food stores and grocery stores in Ljubljana is that both distributions overlap each other in that part of the dispersion line between random and perfectly clustered dispersion.

---

[11] We define a distribution, f_A say, to be more clustered than another, f_B say, if, ceteris paribus, the expected number of empty cells for f_A is greater than that for f_B (see, for example, Table 4.2). If both have the same expected number of empty cells, then f_A is more clustered than f_B if the number of cells with only one member for f_A is greater than that for f_B, and so on.

Finally, we may locate the five classes of retail establishments on the dispersion line, as follows:



The clustered spatial pattern exhibited by food stores and grocery stores at first appears to contradict the common hypothesis in studies of retail structure that convenience goods stores tend to repel one another. However, it is generally accepted that convenience goods stores are strongly influenced in their locational decisions by the spatial distribution of population in an urban area. Thus, if this influencing factor were spatially clustered, its effect on the locational pattern of convenience goods stores would be to cluster them in spite of their hypothesized repelling effect on one another. That is, the apparent spatial clustering of convenience goods outlets may be due to the clustered distribution of population in urban areas and not the result of a spatial affinity between rival establishments. To investigate this possibility, a crude means of measuring the spatial clustering of population in the study area may be devised utilizing data provided by Mrs. Vasle of the Town Planning Institute of Slovenia.

The data on population distribution were broken down by class intervals with 500 persons in each interval. Thus, the logical transfer of this

data into a spatial pattern was to consider every point within the pattern as representing 500 people. This is very crude, to be sure, but it is the only way of utilizing the available information. The resulting spatial dispersion of the residential population in the Ljubljana study area was then fitted by the same methods as were all the other observed frequency distributions of retail establishments. The results are summarized in Tables 4.3 and 4.4.

TABLE 4.3--THE DISTRIBUTION OF THE RESIDENTIAL POPULATION OVER THE 250-METER SQUARES, PARAMETER ESTIMATES AND THE GOODNESS OF FIT

Size of population per square (number of persons)	Number of "500 person" units per square	Number of squares with the indicated size of residential population
0- 500	0	90
501-1,000	1	28
1,001-1,500	2	15
1,501-2,000	3	9
2,001-2,500	4+	2

Model	$u_1 = .6458$ $u_2 = .9856$ $u_2/u_1 = 1.5261$		X^2	d.f.
Binomial	$n = 4$	$p = .1615$	21.69*	1
Poisson	$m = .6458$		13.60*	1
Negative Binomial	$m = .6458$	$k = .9118$	1.59	1
Neyman Type A	$m = .9521$	$a = .6783$	0.31	1

* = X^2 statistic significant at 5 per cent level

On the basis of this crude analysis it appears that the residential population in the study area is highly clustered. This is suggested by the close fit provided by both the negative binomial and the Neyman Type A distributions. Thus there exists considerable justification for retaining

TABLE 4.4--OBSERVED AND EXPECTED DISTRIBUTIONS OF POPULATION IN LJUBLJANA, YUGOSLAVIA

No. of units per cell*	No. of cells observed	Expected Frequency B.	P.	N.B.	T.A.
0	90	71.20	75.49	88.37	90.10
1	28	54.83	48.75	33.41	29.53
2	15	15.84	15.74	13.24	14.85
3	9	2.03	3.39	5.33	6.08
4+	2	.10	.63	3.60	3.45
Total no. of cells =	144	X^2 = 21.69	13.60	1.59	0.31
Total no. of units =	93	$P_{.05}$ = 3.84	3.84	3.84	3.84

* Each unit represents 500 people.

the original hypothesis concerning the agglomerative and deglomerative tendencies of shopping and convenience goods establishments. However, the implication that emerges from the above discussion is that any analysis of the mutual repulsion of convenience goods outlets must separate out the effect of population attraction. In other words, it appears that convenience goods retailers are under the influence of at least two significant locational forces. Because of the unspecialized and recurrent function they perform, such producers tend to be drawn toward the consumers they serve and thus to assume a distributional pattern similar to that of the residential population. At the same time, however, the "locational relation among producers competing for markets is generally one of mutual repulsion represented by the efforts of each seller to find a market where there is not too much

competition."[12] Thus, to the extent that these locational considerations are averaged out with the more dominant consideration counteracting and outweighing the effects of the other. Hoover (1948) suggests a particularly useful analogy of this process by citing methods used in the making of sandpaper "in which an electric charge is imparted to the abrasive particles and an opposite charge to the adhesive-coated paper. The particles are individually attracted to the paper but repelled by each other. The result is that they distribute themselves over the paper in an exceedingly uniform pattern."[13] Hence, the influence of the spatial distribution of purchasing power must be treated separately, but not independently, from the repelling influence of rival establishments. This aspect is taken up in Rogers and Martin (1969), where bivariate quadrat distributions are used to treat both variables simultaneously.

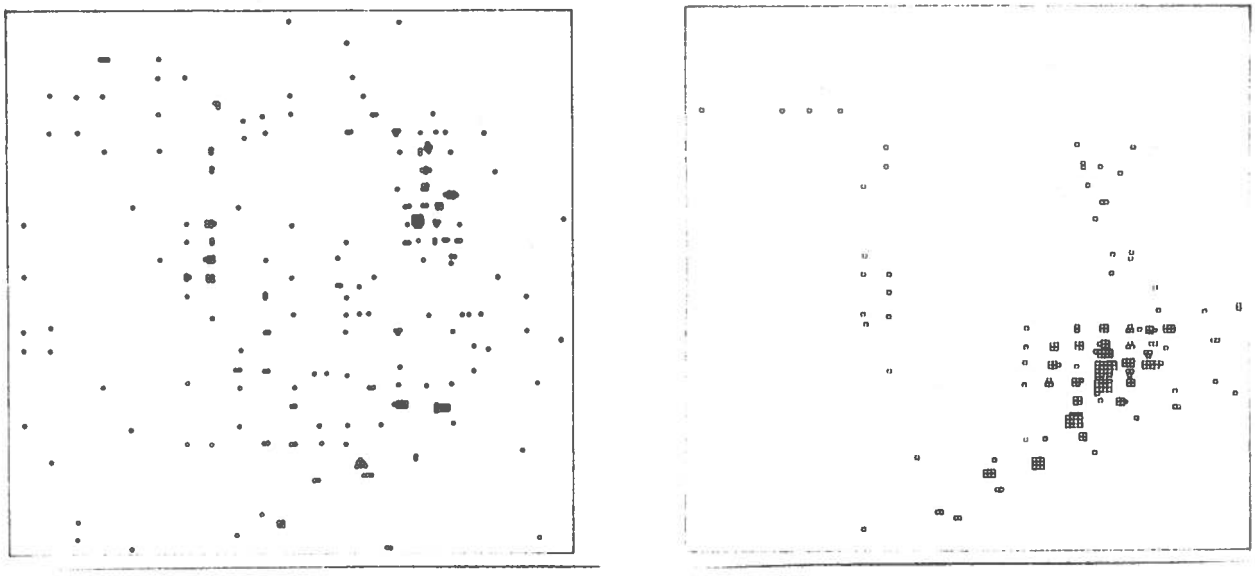
4.3.2 Interurban Comparisons

In order to compare the spatial dispersion of retail establishments in Ljubljana with comparable data for another study area, let us turn to the data collected by the State Department of Employment in California for San Francisco. The data are not exactly comparable for two reasons: (1) the classification system used by the Department of Employment (the S.I.C.) was not identical to the one used in Yugoslavia, and (2) the location of establishments in San Francisco was referenced by block numbers and not by coordinates. Hence, of necessity, we have assigned to all establishments in a single block the coordinates of that block's "center of gravity." In Figure 4.3, however, the establishments are located side by side, and

---

[12] Hoover (1948), p. 48.

[13] Ibid., p. 48, n. 2.



A. FOOD STORES IN SAN FRANCISCO

B. CLOTHING (APPAREL) STORES IN SAN FRANCISCO

FIG. 4.3--THE SPATIAL PATTERNS OF FOOD AND CLOTHING STORES IN SAN FRANCISCO, CALIFORNIA

not on top of one another, for purposes of illustration. The study area, once again, is a square 3,000 by 3,000 meters in size that is gridded into 144 square quadrats, each measuring 250 meters on a side. As in Ljubljana, the study area is centered on the Central Business District. Table 4.5 presents the fits of four distributions to the data on the spatial dispersion of San Francisco's 251 food stores and 221 clothing stores, and Table 4.6 provides a complete list of observed and expected frequencies for food stores.

TABLE 4.5--PARAMETER ESTIMATES AND THE GOODNESS OF FIT FOR FOOD AND CLOTHING ESTABLISHMENTS IN SAN FRANCISCO, CALIFORNIA

a. Food Stores

Model	$u_1 = 1.7431$ $u_2 = 8.4860$ $u_2/u_1 = 4.8684$		X^2	d.f.
Binomial	n = 16	p = .1089	98.00*	3
Poisson	m = 1.7431		79.12*	3
Negative Binomial	m = 1.7431	k = .5988	14.09*	4
Neyman Type A	m = 1.1822	a = 1.4745	28.25*	4

b. Clothing (Apparel) Stores

Model	$u_1 = 1.5347$ $u_2 = 27.4394$ $u_2/u_1 = 17.8790$		X^2	d.f.
Binomial	n = 37	p = .0415	267.28*	3
Poisson	m = 1.5347		253.83*	3
Negative Binomial	m = 1.5347	k = .1024	3.99	3
Neyman Type A	m = .3896	a = 3.9397	32.48*	4

* = X^2 statistic significant at 5 per cent level.

TABLE 4.6--OBSERVED AND EXPECTED DISTRIBUTIONS OF FOOD STORES IN SAN FRANCISCO, CALIFORNIA

No. of stores per cell	No. of cells observed	Expected Frequency B.	P.	N.B.	T.A.
0	59	22.74	25.20	63.64	57.87
1	43	44.49	43.92	28.36	23.09
2	11	40.80	38.28	16.88	21.63
3	8	23.28	22.24	10.88	15.77
4	10	9.25	9.69	7.29	10.34
5	3	2.71	3.38	4.99	6.45
6	1	.61	.98	3.46	3.88
7	1	.11	.24	2.43	2.25
8	0	.01	.05	1.72	1.27
9	3		.01	1.22	.69
10+	5			3.13	.75
Total no. of cells = 144	X^2 = 98.00		79.12	14.09	28.25
Total no. of stores = 251	$P_{.05}$ = 7.82		7.82	9.50	9.50

As in Ljubljana, the spatial dispersion of clothing stores in San Francisco can be satisfactorily described by the negative binomial distribution. Moreover, it appears that the degree of clustering of these stores is roughly the same in the two study areas. For although the mean number of stores per quadrat in San Francisco is about three times as large as the

corresponding figure for Ljubljana, the estimates of the exponent k are almost identical:[14]

Ljubljana	San Francisco
(68 stores)	(221 stores)
m = .4722	m = 1.5347
k = .1038	k = .1024

Observe, however, that unlike the data for Ljubljana, the spatial dispersion of food stores in San Francisco is not satisfactorily accounted for by any of the quadrat distributions that were used for the Ljubljana data (including the Poisson-binomial and the Poisson-negative binomial). It appears that we need a distribution that has more cells containing one store, while at the same time having a large number of empty quadrats (Table 4.6). This suggests that a combination of the negative binomial and the positive binomial, such as the negative binomial-binomial generalized distribution (negative binomial $\vee$ binomial) might provide a satisfactory fit.

---

[14] It is of interest to compare these figures with comparable ones found for Artle's (1965) data on women's clothing stores in downtown Stockholm [Rogers (1965), and Rogers and Gomar (1969)]. For the same size of quadrat, we found an m = 1.0810 and a k = .3050. The quadrat sampling scheme used by Artle, however, differed slightly from the one used in this study. Thus the results are not exactly comparable.

PART V

CONCLUSION

Two fundamental ideas have achieved a more or less central position in theoretical research on intraurban retail spatial structure. These are the idea of order and the idea of causes acting to promote this order. This paper has been concerned largely with a third idea which, since the turn of the century, has become increasingly important in scientific research: the idea of chance.

The many theoretical attempts to account for the spatial disposition of retail activity in urban areas typically assume that such patterns have an identifiable order, an observable regularity which may be generalized to include cities of different sizes, types, and age. This spatial order is, in a sense, imposed on urban patterns by means of a classification system that groups the various uses of land into activity categories, not of identical activities but of activities which seem to behave alike, and then records their occurrence in urban space. The process of ordering is thus an empirical activity involving trial and error. Moreover, the arrangement of activities into groups is predicated on attributes which are judged to be particularly relevant in terms of the kind of problem that is being studied. The only test of the "correctness" of any particular order is its success in accounting for empirical observations; that is, a classification system cannot be judged in advance.

The concept of cause has been a central idea of science since the time of Newton. Thus, it is not surprising to find its appearance in theoretical speculations on retail spatial structure. In effect, the idea of cause in the latter problem context asserts that a definite configuration of factors

acting together in a specified environment will always produce the same observable spatial pattern. The present state of the urban field determines all future states, and thus from a given beginning will always follow the same ends. In short, such a view identifies spatial structure as a machine which possesses definite properties that can be isolated and reproduced in space and time, and the behavior of which can be predicted.

The introduction of the concept of chance does not reject the notion of spatial structure as a machine, but merely asserts that more than one result may occur from the same beginning. Moreover, it specifies that these results must occur in relatively fixed proportions over repeated trials. The idea of order is retained; however, its character now is probabilistic.

Order, cause, and chance were fused in this study to analyze retail spatial patterns. Spatial order was imposed on the geographical arrangement of retail activities in urban areas by grouping them into classes and by describing the spatial distribution of each class by means of a probability function. Cause was introduced by a fundamental hypothesis which held that due to profit considerations, shopping goods stores exhibited a mutual affinity, whereas convenience goods stores, in contrast, tended to repel one another. Finally, retail spatial patterns were analyzed using stochastic models in an attempt to identify the particular chance mechanism which appeared to be producing the observed pattern of spatial dispersion.

More specifically, this study has endeavored to develop means whereby the spatial dispersion of various classes of shopping goods and convenience goods establishments can be quantitatively expressed, systematically analyzed, and compared on both an intraurban and an interurban level. The analysis first translated the study's fundamental hypothesis regarding clustering and dispersal into model form; then deduced the spatial implications of these

assumptions; and finally, empirically tested the deduced implications.

Several implications emerge out of the conceptual, analytical and empirical portions of this study. These refer principally to the description, analysis, and sampling of intraurban retail spatial dispersions.

Description. One of the principal contributions that may be expected from the use of methods such as have been outlined in this study, is the more explicit and effective description of the dispersion of retail establishments in urban areas. It becomes possible in the case of clothing stores, for example, to pose questions concerning the following attributes of their spatial dispersion:

1. the mean number of clothing stores per cell (or cluster) and the corresponding standard error,

2. the mean number of store clusters per cell and the corresponding standard error,

3. the probability of finding r clothing stores in a cell (or cluster),

4. the conditional probability of finding r clothing stores in a cell (or cluster) given that y stores already exist there,

5. the degree of spatial clustering exhibited by clothing stores as measured by their index of clustering k,[15]

6. the position of k for clothing stores with respect to other classes of shopping goods (intraclass comparisons),

7. the change of k in time (temporal comparisons),

8. the change of k in space (interurban comparisons).

---

[15] It should be noted that attributes 5, 6, 7, and 8 apply only to patterns described by the negative binomial distribution.

It has been suggested that the paucity of empirical data on the physical configurations of urban populations and activities is one of the principal impediments to the development of a viable theory of intraurban spatial structure [Jones (1960)]. The need for appropriate measures with which to condense and summarize such data into comprehensible form is clear. The methods outlined in this study seem to have considerable merit in this respect. They may be used to answer questions concerning the temporal changes of a city's retail spatial pattern. They indicate the mechanisms which appear to generate this structure. Finally, they allow interurban comparisons to be made, a fundamental requisite in theory building.

Analysis. The analytic portions of this study sought to establish the chance mechanisms which produce the observed spatial dispersion of retail establishments in urban areas. The procedure that was followed in each instance was to hypothesize a particular behavioral tendency on the part of a given class of retailers and to deduce from this assumption a spatial distribution which then could be subjected to empirical testing. The implication that evolves from the principal findings described above is a qualified support of the fundamental hypothesis of this study: that ceteris paribus rival shopping goods producers tend to attract one another and thus to cluster, whereas competing convenience goods producers tend to repel one another and thus tend to assume a less clustered spatial pattern. The latter conclusion, however, must be qualified to allow for the clustering of convenience goods firms in instances of a spatially clustered population.

The distributions which were used to describe retail spatial dispersions were derived from a set of initial propositions. Thus, in addition to their utility as a descriptive medium, they serve an analytic purpose as well. That is, the distributions evolved out of specific assumptions concerning

the spatial behavior of retailers. Therefore, in instances where a particular theoretical distribution provides a satisfactory fit to observed data, the model which generated this distribution immediately suggests itself as the chance mechanism responsible for the corresponding empirical distribution. This feature of models derived by stochastic processes is analogous to the latter's deterministic counterpart in the physical and natural sciences--differential equations.

Spatial Sampling. Spatial sampling is the process of taking samples over a spatial domain when some information is available regarding the geographical arrangement of the members of the population that is being sampled in this domain. The need for quantitative records of spatially distributed phenomena in urban areas has generated a growing interest in methods of spatial sampling. Frequently, the members of an entire population cannot be enumerated because of time or cost constraints. And it becomes important, in such instances, to determine what constitutes an adequate sample design.

Quadrat analysis can provide useful information for spatial sampling. Consider, for example, the problem of designing an on-going survey to produce estimates of the monthly retail employment in a city and the distribution of this employment by the different major categories of retail trade. Moreover, assume that the spatial dispersions of the various categories of this employment are known. An efficient sample design would take advantage of this information. Thus, for example, randomly dispersed employment would be sampled differently than negative binomially dispersed employment.

In sum, it appears that three sets of implications are indicated by the results of this study. These suggest that quadrat analysis may be of considerable use in future efforts to describe, analyze and sample retail spatial dispersion in urban areas. However, this conclusion is very tentative and awaits the results of future empirical and theoretical efforts for its verification or rejection.

BIBLIOGRAPHY

1. ADELSON, R. M. (1966). Compound Poisson distributions. Operational Research Quarterly, 17, 73-75.

2. ANSCOMBE, F. J. (1948). The transformation of Poisson, binomial and negative binomial data. Biometrika, 35, 246-254.

3. ANSCOMBE, F. J. (1949). The statistical analysis of insect counts based on the negative binomial distribution. Biometrics, 5, 165-173.

4. ANSCOMBE, F. J. (1950). Sampling theory of the negative-binomial and logarithmic series distributions. Biometrika, 37, 358-382.

5. ARBOUS, A. G. and KERRICH, J. C. (1951). Accident statistics and the concept of accident-proneness. Biometrics, 7, 340-432.

6. ARCHIBALD, E. E. A. (1948). Plant populations. I. A new application of Neyman's contagious distribution. Annals of Botany, London, N.S., 12, 221-235.

7. ARCHIBALD, E. E. A. (1950). Plant populations. II. The estimation of the number of individuals per unit area of species in heterogeneous plant populations. Annals of Botany, London, N.S., 12, 221-235.

8. ARTLE, R. K. (1965). The Structure of the Stockholm Economy, American edition. Cornell University Press, Ithaca, New York.

9. AUBERT-KRIER, J. (1954). Monopolistic and imperfect competition in retail trade, in Monopoly and Competition and their Regulation, ed., E. H. Chamberlin. Macmillan, London, pp. 281-300.

10. BATES, G. E. and NEYMAN, J. (1952). Contributions to the theory of accident proneness. I. An optimistic model of the correlation between light and severe accidents. University of California Publication in Statistics, 1, pp. 215-253.

11. BEALL, G. (1940). The fit and significance of contagious distributions when applied to observations on larval insects. Ecology, 21, 460-474.

12. BEALL, G. and RESCIA, R. R. (1953). A generalization of Neyman's contagious distributions. Biometrics, 8, 334-386.

13. BERKSON, J. (1938). Some difficulties of interpretation encountered in the application of the chi-square test. Journal of the American Statistical Association, 33, 526-536.

14. BERKSON, J. (1940). A note on the chi-square test, the Poisson and the binomial. Journal of the American Statistical Association, 35, 362-367.

15. BERRY, B. J. L. and GARRISON, W. L. (1958a). A note on central place theory and the range of a good. Economic Geography, 34, 304-311.

16. BERRY, B. J. L. and GARRISON, W. L. (1958b). Functional bases of the central place hierarchy. Economic Geography, 34, 145-154.

17. BHATTACHARYA, S. K. and HOLLA, M. S. (1965). On a discrete distribution with special reference to the theory of accident proneness. Journal of the American Statistical Association, 60, 1060-1066.

18. BLACKMAN, G. E. (1942). Statistical and ecological studies in the distribution of species in plant communities. I. Dispersion as a factor in the study of changes in plant populations. Annals of Botany, London, N.S., 6, 351-370.

19. BLISS, C. I. and FISHER, R. A. (1953). Fitting the negative binomial distribution to biological data, and note on the efficient fitting of the negative binomial. Biometrics, 9, 176-200.

20. BLISS, C. I. and OWEN, A. R. G. (1958). Negative binomial distributions with a common k. Biometrika, 45, 37-58.

21. BUNGE, W. (1962). Theoretical Geography. C. W. K. Gleerup, Lund, Sweden.

22. CAMP, B. H. (1940). Further comments on Berkson's problem. Journal of the American Statistical Association, 35, 368-376.

23. CAMPBELL, J. T. (1934). The Poisson correlation function. Proceedings of the Edinburgh Mathematical Society, Ser. 2, 4, 18-26.

24. CHERNOFF, H. and LEHMANN, E. L. (1954). The use of maximum likelihood estimates in X^2 tests for goodness of fit. Annals of Mathematical Statistics, 25, 579-586.

25. CLARK, P. J. (1956). Grouping in spatial distributions. Science, 123, 373, 374.

26. CLARK, P. J. and EVANS, F. C. (1954). Distance to nearest neighbor as a measure of spatial relationships in populations. Ecology, 35, 445-453.

27. CLARK, P. J. and EVANS, F. C. (1954). On some aspects of spatial pattern in biological populations. Science, 121, 397-398.

28. COCHRAN, W. G. (1952). The X^2 test of goodness of fit. Annals of Mathematical Statistics, 23, 315-345.

29. COCHRAN, W. G. (1954). Some methods for strengthening the common X^2 tests. Biometrics, 10, 417-451.

30. COCHRAN, W. G. (1936). The X^2 distribution for the binomial and Poisson series, with small expectations. Annals of Eugenics, 7, 207-217.

31. CURTIS, J. T. and McINTOSH, R. P. (1950). The interrelations of certain analytic and synthetic phytosociological characters. Ecology, 31, 434-455.

32. DACEY, M. F. (1962). Analysis of central place and point patterns by a nearest neighbor method. Proceedings of the IGU Symposium in Urban Geography, Lund 1960, Ed. Kurt Norborg, Lund Studies in Geography, B, 24, 55-75.

33. DACEY, M. F. (1963). Order neighbor statistics for a class of random patterns in multidimensional space. Annals of the Association of American Geographers, 53, 505-515.

34. DACEY, M. F. (1965). The geometry of central place theory. Geografiska Annaler, B, 47, 111-124.

35. DALENIUS, T., HAJEK, J., and ZUBRZYCKI, S. (1961). On plane sampling and related geometrical problems. Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics, and Probability, University of California Press, Berkeley, pp. 125-150.

36. DANIELS, H. E. (1961). Mixtures of geometric distributions. Journal of Royal Statistical Society, B, 23, 409-413.

37. DOUGLAS, J. B. (1955). Fitting the Neyman Type A (two parameter) contagious distribution. Biometrics, 11, 149-173.

38. DUBEY, S. D. (1966). Compound Pascal distributions. Annals of the Institute of Statistical Mathematics, Tokyo, 18, 357-365.

39. EDWARDS, C. B. and GURLAND, J. (1961). A class of distributions applicable to accidents. Journal of the American Statistical Association, 56, 503-517.

40. EVANS, D. A. (1953). Experimental evidence concerning contagious distributions in ecology. Biometrika, 40, 186-211.

41. FELLER, W. (1943). On a general class of "contagious" distributions. Annals of Mathematical Statistics, 14, 389-400.

42. FELLER, W. (1957). An Introduction to Probability Theory and Its Application. John Wiley, New York.

43. FISHER, R. A. (1924). The conditions under which the chi square measures the discrepancy between observation and hypothesis. Journal of the Royal Statistical Society, 87, 442-450.

44. FISHER, R. A. (1941). The negative binomial distribution. Annals of Eugenics, 11, 182-187.

45. FISHER, R. A., CORBET, A. S. and WILLIAMS, C. B. (1943). The relation between the number of species and the number of individuals in a random sample from an animal population. Journal of Animal Ecology, 12, 42-58.

46. FIX, E. (1949). Tables of the noncentral χ^2, University of California Press, Berkeley.

47. GETIS, A. (1964). Temporal land use pattern analyses with the use of the nearest neighbor and quadrat methods. Annals of the Association of American Geographers, 54, 391-398.

48. GREENWOOD, H. and YULE, G. U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings. Journal of the Royal Statistical Society, 83, 255-279.

49. GREIG-SMITH, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. Annals of Botany, London, N.S., 16, 293-316.

50. GREIG-SMITH, P. (1964). Quantitative Plant Ecology. 2nd Ed. Butterworths, London.

51. GURLAND, J. (1957). Some interrelations among compound and generalized distributions. Biometrika, 44, 265-268.

52. GURLAND, J. (1958). A generalized class of contagious distributions. Biometrics, 14, 229-249.

53. GURLAND, J. (1959). Some applications of the negative binomial and other contagious distributions. American Journal of Public Health, 49, 1388-1399.

54. HAIGHT, F. A. (1959). The generalized Poisson distribution. Annals of the Institute of Statistical Mathematics, Tokyo, 11, 101-105.

55. HAIGHT, F. A. (1967). Handbook of the Poisson Distribution. John Wiley, New York.

56. HALDANE, J. B. S. (1941). The fitting of binomial distributions. Annals of Eugenics, 11, 179-181.

57. HODGES, J. L., Jr. and Le CAM, L. (1960). The Poisson approximation to the Poisson binomial distribution. Annals of Mathematical Statistics, 31, 737-740.

58. HOLGATE, P. (1964). Estimation for the bivariate Poisson distribution. Biometrika, 51, 241-245.

59. HOLGATE, P. (1965a). Tests of randomness based on distance methods. Biometrika, 52, 345-353.

60. HOLGATE, P. (1965b). The distance from a random point to the nearest point of a closely packed lattice. Biometrika, 52, 261-263.

61. HOLGATE, P. (1966). Bivariate generalizations of Neyman's Type A distribution. Biometrika, 53, 241-244.

62. HOLLA, M. S. and BHATTACHARYA, S. K. (1964). On a discrete compound distribution. Annals of the Institute of Statistical Mathematics, Tokyo, 16, 377-384.

63. HOTELLING, H. (1929). Stability in competition. Economic Journal, 39, 41-57.

64. HOOVER, E. M. (1948). The Location of Economic Activity, McGraw-Hill, New York.

65. HUDSON, J. C. and FOWLER, P. M. (1965). The Concept of Pattern in Geography, Discussion Paper No. 1. Department of Geography, University of Iowa, Iowa City.

66. ISHII, G. and HAYAKAWA, R. (1960). On the compound binomial distribution. Annals of the Institute of Statistical Mathematics, Tokyo, 12, 69-80.

67. JONES, B. G. (1960). The theory of the urban economy: origins and development with emphasis on intraurban distribution of population and economic activity. Ph.D. thesis, University of North Carolina, Chapel Hill, North Carolina.

68. KATTI, S. K. (1960). Some aspects of statistical inference for contagious distributions. Ph.D. thesis, Iowa State University, Ames, Iowa.

69. KATTI. S. K. and GURLAND, J. (1961). The Poisson Pascal distribution. Biometrics, 17, 527-538.

70. KATTI, S. K. and GURLAND, J. (1962a). Some methods of estimation for the Poisson binomial distribution. Biometrics, 18, 42-51.

71. KATTI, S. K. and GURLAND, J. (1962b). Efficiency of certain methods of estimation for the negative binomial and the Neyman Type A distributions. Biometrika, 49, 215-226.

72. KEMP, C. D. and KEMP. A. W. (1956). The analysis of point quadrat data. Australian Journal of Botany, 4, 167-174.

73. KEMP, C. D. (1967). On a contagious distribution suggested for accident data. Biometrics, 23, 241-255.

74. KENDALL, D. G. (1948). On some modes of population growth leading to R. A. Fisher's logarithmic series distribution. Biometrika, 35, 6-15.

75. KENDALL, M. G. and STUART, A. (1961) The Advanced Theory of Statistics, II: Inference and Relationship. Charles Griffin, London.

76. KHATRI, C. G. and PATEL, I. R. (1961). Three classes of univariate discrete distributions. Biometrics, 17, 567-575.

77. KHATRI, C. G. (1962). A fitting procedure for a generalized binomial distribution. Annals of the Institute of Statistical Mathematics, Tokyo, 14, 133-141.

78. KRISHNAMOORTHY, A. S. (1951). Multivariate binomial and Poisson distribution. Sankhya, 11, 117-124.

79. LEWIS, W. A. (1945). Competition in retail trade, Economica, 12, 202-234.

80. MCCONNELL, H. (1966). Quadrat Methods in Map Analysis, Discussion Paper No. 3. Department of Geography, University of Iowa, Iowa City.

81. MCGUIRE, J. U., BRINDLEY, T. A. and BANCROFT, T. A. (1957). The distribution of European corn borer larvae pyrausta nubilalis (hbn) in field corn. Biometrics, 13, 65-78; 14, 432-434 (errata and extensions).

82. MACEDA, E. C. (1948). On the compound and generalized Poisson distributions. Annals of Mathematical Statistics, 19, 414-416.

83. MANN, H. B. and WALD, A. (1942). On the choice of the number of intervals in the application of the X^2 test. Annals of Mathematical Statistics, 13.

84. MARITZ, J. S. (1950). On the validity of inferences drawn from the fitting of Poisson and negative binomial distributions to observed accident data. Psychological Bulletin, 47, 434-443.

85. MARITZ, J. S. (1952). Note on a certain family of discrete distributions. Biometrika, 39, 196-198.

86. MARTIN, D. C. and KATTI, S. K. (1965). Fitting of certain contagious distributions to some available data by the maximum likelihood method. Biometrics, 21, 34-48.

87. MARTIN, J., GOMAR, N. and ROGERS, A. (1969). Some Computer Programs for Quadrat Analysis, Working Paper. Center for Planning and Development Research, University of California, Berkeley.

88. MATUSZEWSKI, T. I. (1962). Some properties of Pascal distribution for finite population. Journal of the American Statistical Association, 57, 172-174.

89. MOORE, P. J. (1954). Spacing in plant populations. Ecology, 35, 222-227.

90. MORISITA, M. (1954). Estimation of population density by spacing method. Memoirs of the Faculty of Science, Kyushu University, E, I, 187-197.

91. NAUS, J. I. (1965). Clustering of random points in two dimensions. Biometrika, 52, 263-267.

92. NELSON, R. L. (1958). The Selection of Retail Location . F. W. Dodge, New York.

93. NEYMAN, J. (1939). On a new class of "contagious" distributions, applicable in entomology and bacteriology. Annals of Mathematical Statistics, 10, 35-37.

94. NEYMAN, J. and SCOTT, E. L. (1952). A theory of the spatial distribution of galaxies. Astrophysical Journal, 116, 144-163.

95. NEYMAN, J. and SCOTT, E. L. (1957). On a mathematical theory of populations conceived as conglomerations of clusters. Cold Spring Harbour Symposia on Quantitative Biology, 22, 109-120.

96. NEYMAN, J. and SCOTT, E. L. (1959). Stochastic models of population dynamics. Science, 130, 303-308.

97. O'CARROLL, F. M. (1962). Fitting a negative binomial distribution to coarsely grouped data by maximum likelihood. Applied Statistics, 11, 196-201.

98. OTTESTAD, P. (1944). On certain compound frequency distributions. Skandinavisk Aktuarietidskrift, 27, 32-42.

99. PARZEN, E. (1960). Modern Probability Theory and Its Applications. John Wiley, New York.

100. PATNAIK, P. B. (1949). The non-central χ^2- and F-distributions and their applications. Biometrika, 36, 202-232.

101. PATIL, G. P. (1960). On the evaluation of the negative binomial distribution with examples. Technometrics, 2, 501-505.

102. PATIL, G. P. (1962). Maximum likelihood estimation for generalized power series distributions and its application to a truncated binomial distribution. Biometrika, 49, 227-238.

103. PATIL, G. P. (1962). Certain properties of the generalized power series distribution. Annals of the Institute of Statistical Mathematics, Tokyo, 14, 179-182.

104. PATIL, G. P. (1964). On certain compound Poisson and compound binomial distributions. Sankhyā, A, 26, 293-294.

105. PATIL, G. P. (1965). Certain characteristic properties of multivariate discrete probability distributions akin to the Bates-Neyman model in the theory of accident proneness. Sankhyā, A, 27, 259-270.

106. PATIL, G. P. and BILDIKAR, S. (1967). Multivariate logarithmic series distribution as a probability model in population and community ecology and some of its statistical properties. Journal of the American Statistical Association, 62, 655-674.

107. PIELOU, E. C. (1957). The effect of quadrat size on the estimation of the parameters of Neyman's and Thomas' distributions. Journal of Ecology, 45, 31-47.

108. PIELOU, E. C. (1960). A single mechanism to account for regular, random and aggregated populations. Journal of Ecology, 48, 575-584.

109. QUENOUILLE, M. H. (1949). A relation between the logarithmic, Poisson, and negative binomial series. Biometrics, 5, 162-164.

110. RATCLIFF, R. U. (1939). The Problem of Retail Site Selection. Bureau of Business Research, School of Business Administration, University of Michigan, Ann Arbor, Michigan.

111. ROBINSON, P. (1954). The distribution of plant populations. Annals of Botany, London, N.S., 18, 35-45.

112. ROGERS, A. (1965). A stochastic analysis of the spatial clustering of retail establishments. Journal of the American Statistical Association, 60, 1094-1103.

113. ROGERS, A. and GOMAR, N. (1969). Statistical inference in quadrat analysis. Geographical Analysis, forthcoming.

114. ROGERS, A. and MARTIN, J. (1969). Bivariate Quadrat Analysis of the Dispersion of Employment and Population in Urban Areas, Working Paper. Center for Planning and Development Research, University of California, Berkeley.

115. SATTERTHWAITE, F. E. (1942). Generalized Poisson distribution. Annals of Mathematical Statistics, 13, 410-417.

116. SCHILLING, W. (1947). A frequency distribution represented as the sum of two Poisson distributions. Journal of the American Statistical Association, 42, 407-424.

117. SHENTON, L. R. (1949). On the efficiency of the method of moments and Neyman's Type A distribution. Biometrika, 36, 450-454.

118. SHENTON, L. R. (1950). Maximum likelihood and the efficiency of the method of moments. Biometrika, 37, 111-116.

119. SHENTON, L. R. (1958). Moment estimators and maximum likelihood. Biometrika, 45, 411-420.

120. SHENTON, L. R. (1959). The distribution of moment estimators. Biometrika, 46, 296-305.

121. SHENTON, L. R. and WALLINGTON, P. A. (1962). The bias of moment estimators with an application to the negative binomial distribution. Biometrika, 49, 193-204.

122. SHUMWAY, R. and GURLAND, J. (1960a). Fitting the Poisson binomial distribution. Biometrics, 16, 522-533.

123. SHUMWAY, R. and GURLAND, J. (1960b). A fitting procedure for some generalized Poisson distributions. Skandinavisk Aktuareitidskrift, 43, 87-108.

124. SIBUYA, M., YOSHIMURA, I. and SHIMIZU, R. (1964). Negative multinomial distribution. Annals of the Institute of Statistical Mathematics, 16, 409-426.

125. SICHEL, H. S. (1951). The estimation of the parameters of a negative binomial distribution with special reference to psychological data. Psychometrika, 16, 107-127.

126. SKELLAM, J. G. (1948). A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. Journal of the Royal Statistical Society, B, 10B, 257-261.

127. SKELLAM, J. G. (1952). Studies in statistical ecology. Biometrika, 39, 346-362.

128. SKELLAM, J. G. (1958). On the derivation and applicability of Neyman's Type A distribution. Biometrika, 45, 32-36.

129. SOMERVILLE, P. N. (1957). Optimum sampling in binomial populations. Journal of the American Statistical Association, 52, 494-502.

130. SPROTT, D. A. (1958). The method of maximum likelihood applied to the Poisson binomial distribution. Biometrika, 14, 97-106.

131. SUBRAHMANIAM, K. (1965). On a general class of contagious distributions: The Pascal-Poisson distribution. Paper No. 356. Department of Biostatistics, Johns Hopkins University, Baltimore, Maryland.

132. SUKHATME, P. V. (1938). On the distribution of X^2 in samples of the Poisson series. Journal of the Royal Statistical Society, 1, 75-79.

133. TEICHER, H. (1954). On the multivariate Poisson distribution. Skandinavisk Aktuareitidskrift, 37, 1-9.

134. TEICHER, H. (1960). On the mixture of distribution. Annals of Mathematical Statistics, 31, 55-73.

135. THOMAS, M. (1949). A generalization of Poisson's binomial limit for use in ecology. Biometrika, 36, 18-25.

136. THOMPSON, H. R. (1954). A note on contagious distributions. Biometrika, 41, 268-271.

137. THOMPSON, H. R. (1955). Spatial point processes with application to ecology. Biometrika, 42, 102-115.

138. THOMPSON, H. R. (1956). Distribution of distance to n-th neighbor in a population of randomly distributed individuals. Ecology, 37, 391-394.

139. THOMPSON, H. R. (1958). The statistical study of plant distribution patterns using a grid of quadrats. Australian Journal of Botany, 6, 322-343.

140. TORII, T. (1956). The Stochastic Approach in Field Population Ecology. Japan Society for the Promotion of Science, Tokyo.

141. TWEEDIE, M. C. K. (1952). The estimation of parameters from sequentially sampled data on a discrete distribution. Journal of the Royal Statistical Society, B, 14, 238-245.

142. WALSH, J. E. (1963). Bounded probability properties of Kolmogorov-Smirnov and similar statistics for discrete data. Annals of the Institute of Statistical Mathematics, 15, 153-158.

143. WILLIAMS, C. A., Jr. (1950). On the choice of the number and width of classes for the chi-square test of goodness of fit. Journal of the American Statistical Association, 45, 77-86.

Center for Planning and Development Research
Working Papers Sponsored by the
Economic Development Administration

#65 Ronald Choy, "A Review of the Role of Demographic Variables in Macro-Econometric Models," August, 1967.

#66 Andrei Rogers, "The Aggregation Problem in Demography," October, 1967.

#67 Michael Teitz, "Toward A Theory of Urban Public Facility Location," November, 1967.

#70 Michael Teitz, "Locational Strategies for Competitive Systems," December, 1967.

#72 William Alonso, "The Quality of Data and the Choice and Design of Predictive Models," December, 1967.

#74 William Alonso, "Industrial Location and Regional Policy in Economic Development," February, 1968.

#78 William Alonso, "Equity and Its Relation to Efficiency in Urbanization," July, 1968.

#81 Andrei Rogers and Susan McDougall, "An Analysis of Population Growth and Change in Slovenia and in the Rest of Yugoslavia," June, 1968.

#84 Andrei Rogers, "Activity Allocation Models in Transportation Planning," September, 1968.

#86 Andrei Rogers and Norbert G. Gomar, "Statistical Inference in Quadrat Analysis," October, 1968.

#87 William Alonso, "Aspects of Regional Planning and Theory in the United States," October, 1968.

#89 David Darwent, "Growth Pole and Growth Center Concepts: A Review, Evaluation and Bibliography," October, 1968.

#90 William Alonso, "Beyond the Inter-Disciplinary Approach to Planning," November, 1968.

#91 Shlomo Angel, "Point Patterns in Urban Regions: Two Complementary Tests," January, 1969.

#92 Andrei Rogers and Susan McDougall Choy, "Multisectoral Models of Regional Demographic and Economic Growth," February, 1969.

#93 Andrei Rogers, "Quadrat Analysis of Urban Dispersion," March, 1969.