PART TWO

SPATIAL PATTERN IN ANIMAL AND PLANT POPULATIONS

Organisms are not spread at random across the landscape, and one important question of landscape ecology is the pattern of individuals in space. This question is important in its own right since for zoologists spacing behavior has been a central problem in behavioral ecology and for botanists the investigation of pattern has played a prominent role in the methods for study of plants as individuals.

The investigation of spatial pattern has become a specialized area of study in plant ecology and an imposing array of methods has developed over the last 50 years (Dale 1997). In this section we will introduce only a few of the methods that can be used, and interested students are encouraged to go to the more specialized literature. Similarly animal ecologists have developed a large literature on home range estimation and territoriality in animals that we can only touch on here.

Our primary focus is on spatial pattern from two perspectives. First, we wish to determine what spatial pattern a population shows in the field. The spatial pattern of a population has consequences for the estimation of abundance and the construction of confidence intervals for these estimates. In addition, if we are to census a population efficiently, the design of our sampling program will be affected by the spatial pattern of the organisms under study. These are practical matters addressed toward statistical problems of estimation. Second, we need to develop a set of metrics that will provide a measure of spatial pattern. If one population of shrubs has a more clumped pattern than another population, how can we measure the degree of clumping? The metrics of spatial pattern are usually referred to as *indices of dispersion*, and there is a large literature on how to measure quantitatively spatial pattern, or more directly which

index of dispersion is best. We will approach this problem with a quantitative outlook in this section, but I would hasten to remind you that the important part of spatial pattern is the ecology behind it, the mechanisms that generate and maintain spatial patterns. We will not deal with these mechanisms in this book, but only with the foundation of the metrics of measuring spatial pattern efficiently and accurately.

CHAPTER 6

SPATIAL PATTERN AND INDICES OF DISPERSION

Page

Organisms may random, uniform, or clumped spatial patterns in nature. The patterns we recognize are partly a function of scale of the study area. In a sense all field ecologists recognize this, since looked at on a global scale any tree species must have an aggregated distribution because trees do not occur in the ocean. So begin this chapter with the realization that a given population can have a random distribution at one scale, an aggregated distribution at another scale, and a uniform distribution at

yet another scale. The first question we must ask once we have fixed the scale of the study is how we can decide statistically which of these patterns is an adequate description for our particular population. We have already addressed this question peripherally for the case of quadrat sampling in Chapter 4, and in this chapter we carry this discussion forward with more sophisticated and reliable measures for spatial pattern.

Two situations separate the methods for pattern detection in plant and animal populations in the field. First, we may have a complete spatial map of the population we are studying. This is an ideal situation because we can apply an array of sophisticated mathematical methods to such spatial maps (Diggle 1983). When we have a spatial map we know the density of the population because it is completely enumerated. Spatial maps can be constructed for plants in quadrats, for bird territories in forests or grasslands, or for nest sites of geese. Many hypotheses about ecological processes can be tested with spatial maps, and much effort has been put into obtaining data to construct them for many populations of plants (Losos and Leigh 2004, Cao et al. 2008).

Second, we may be interested in broad scale sampling of populations that cannot be completely mapped and enumerated in order to quantify pattern. Forest ecologists typically deal with extensive stands of trees, and yet need information on spatial pattern both to test ecological hypotheses and to estimate abundance efficiently. Many of the methods we have already discussed in the last 3 chapters can be used to quantify spatial patterns, and in addition an array of new methods based on contiguous quadrats can be utilized to measure spatial pattern.

Figure 6.1 gives a flow diagram from Fortin and Dale (2005) that is most useful for studies of spatial pattern in plant and animal populations because it focuses the mind on what is the question and what are the resulting actions you should take. This chapter can only be a simple introduction to the methods they describe in their book and new spatial methods that are evolving with the computer age.
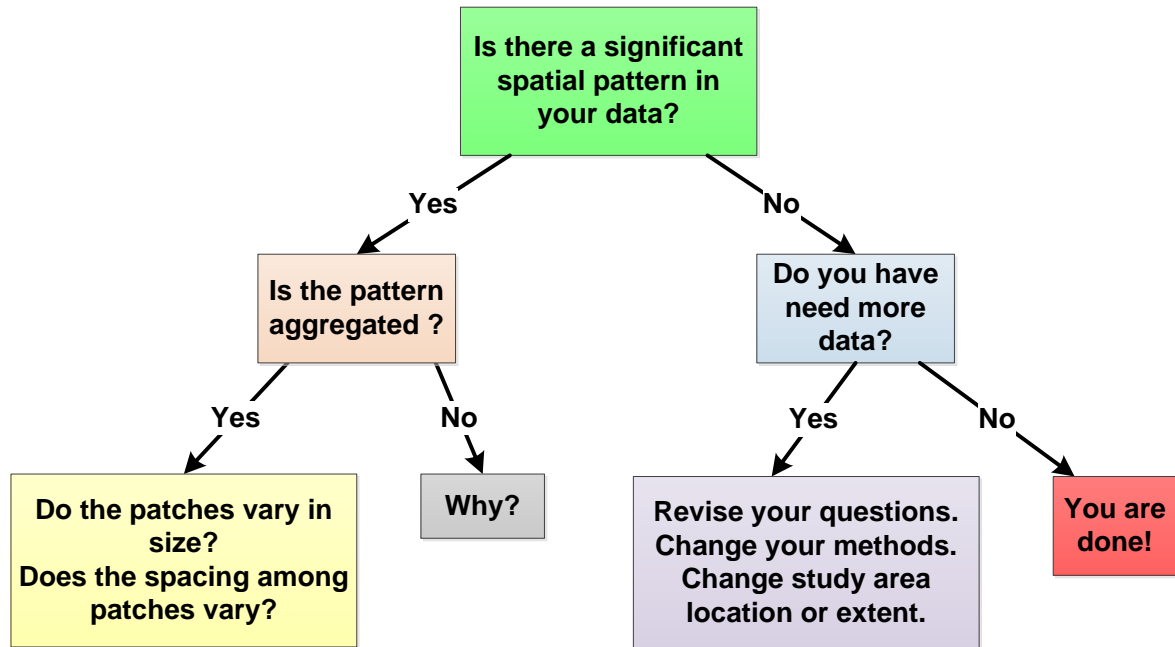
Figure 6.1 Key questions you should ask while analysing spatial ecological data. (From Fortin and Dale 2005.)

6.1 METHODS FOR SPATIAL MAPS

6.1.1 Nearest Neighbor Methods

In some situations an ecologist may have an exact map of the geographical location of each individual organism. Figure 6.2 illustrates this type of spatial data for redwood seedlings in California. Such data represent a complete enumeration of the population being studied, and consequently population density is *known* and not estimated[*].

Spatial maps like Figure 6.2 can be used to measure the pattern of a population.

---

[*] If, on the other hand, there are several quadrats like that shown in Fig. 6.1, one is dealing with quadrat counts - see Chapter 4.

REDWOOD SEEDLINGS
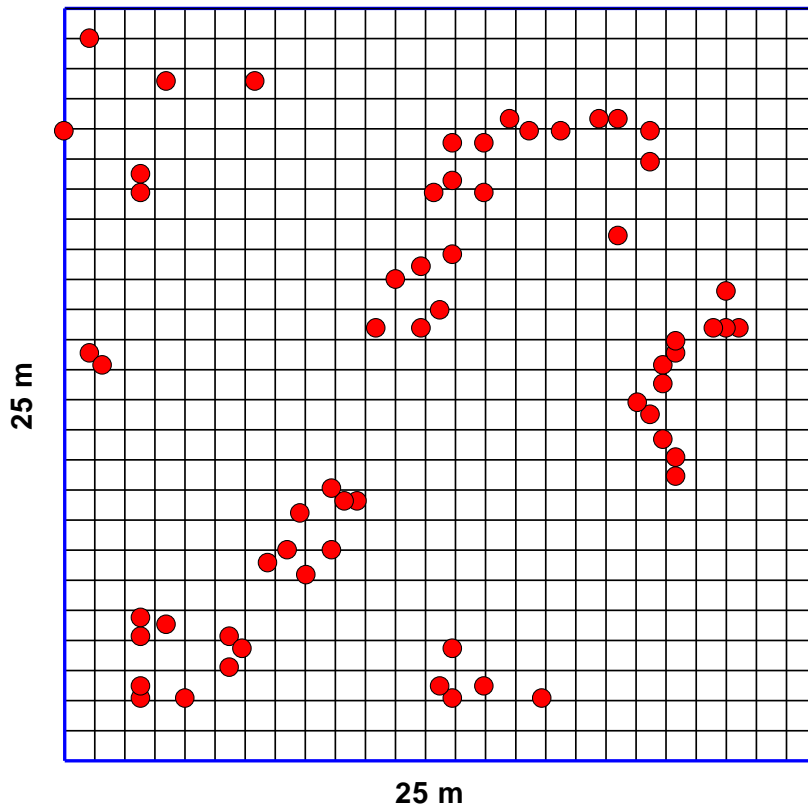
Figure 6.2 Spatial map of the locations of 62 redwood tree seedlings *(Sequoia sempervirens)* in a 25 X 25 m quadrat in California. (Data from Strauss, 1975.)

Clark and Evans (1954) were the first to suggest a method for analyzing pattern for spatial maps. They used the distance from an individual to its nearest neighbor (Figure 5.6, A→ B) as the relevant measure. These distances are measured for *all* individuals in the spatial map, and thus there are no problems with random sampling because it is a complete enumeration of this local population. The nearest neighbor of an individual within the study area can be an individual located outside of the study area (see Fig. 5.1). For this simple situation we have:

$$\bar{r}_A = \text{Mean distance to nearest neighbor} = \frac{\sum r_i}{n}$$

(6.1)

where r_i = Distance to nearest neighbor for individual *i*

   n = Number of individuals in study area

Clark and Evans (1954) showed that the expected distance to the nearest-neighbor could be calculated very simply for a large population that has a random spatial pattern. Define:

$$\rho = \text{Density of organisms} = \frac{\text{Number in study area}}{\text{Size of study area}}$$

then:

$$\bar{r}_E = \text{Expected distance to nearest neighbor} = \frac{1}{2\sqrt{\rho}} \tag{6.2}$$

We can measure the deviation of the observed pattern from the expected random pattern by the ratio:

$$R = \frac{\bar{r}_A}{\bar{r}_E} = \text{Index of aggregation} \tag{6.3}$$

If the spatial pattern is random, $R = 1$. When clumping occurs, R approaches zero and in a regular pattern R approaches an upper limit around 2.15.

A simple test of significance for deviation from randomness is available because the standard error of the expected distance is known exactly from plane geometry:

$$z = \frac{\bar{r}_A - \bar{r}_E}{s_r} \tag{6.4}$$

where z = Standard normal deviate

s_r = Standard error of the expected distance to nearest neighbor

$= 0.26136/\sqrt{n\rho}$

n = Number of individuals in study area

ρ = Density of individuals in study area

Consider an example. Campbell and Clarke (1971) measured nearest neighbor distances for 39 singing crickets (*Teleogryllus commodus*) in Australia. They obtained these data for nearest neighbor distances:

$$n = 39 \qquad \sum r_i = 63.4 \qquad \sum (r_i)^2 = 136.78$$

$$\bar{r}_A = \frac{\sum \bar{r}_i}{n} = \frac{63.4}{39} = 1.626 \text{ meters}$$

For this population of 39 animals in a circular area of 12 m radius (452.4 m^2) the density is 0.08621 crickets per square meter. Thus from equation (6.2):

$$\bar{r}_E = \frac{1}{2\sqrt{0.08621}} = 1.703 \text{ meters}$$

The index of aggregation is:

$$R = \frac{\bar{r}_A}{\bar{r}_E} = \frac{1.626}{1.703} = 0.95$$

which indicates a slight tendency toward an aggregated pattern away from randomness. To test for significant deviation from random pattern we use equation (6.4):

$$z = \frac{\bar{r}_A - \bar{r}_E}{s_r} = \frac{1.626 - 1.703}{0.26136/\sqrt{(39)(0.08621)}} = -0.54$$

Since $|z|$ is less than 1.96, we tentatively accept the null hypothesis at $\alpha = 0.05$ that these crickets are randomly spaced in the study area.

The Clark and Evans test (eq. 6.4) is unbiased only when a boundary strip is included in the study. But in many cases ecologists have not used boundary strips because they wish to maximize sample size (e.g. Fig. 6.2). If the Clark and Evans test is applied in this situation without using a boundary strip, the test is biased. Organisms close to the boundary will tend to have larger nearest-neighbor distances than those well inside (e.g. individual C in Fig. 5.5). Sinclair (1985) showed that the uncorrected Clark and Evans test (eq. 6.4) is biased in favor of regular patterns, so that many aggregated patterns are judged to be random and many random patterns are judged to be uniform. This bias is enormous with small sample sizes ($n < 100$).

If a boundary strip is *not* included in the study, you should use the Donnelly (1978) modification of the Clark and Evans test as follows:

$\bar{r}_c$ = Expected distance to nearest neighbor corrected for lack
        of a boundary strip

$$= \bar{r}_E + \left[\left(0.051 + \frac{0.041}{\sqrt{n}} \right) \left(\frac{L}{n} \right) \right]$$

(6.5)

where L = length of the boundary of the whole study area

The standard error of this expected distance is given by Donnelly as:

$$s_r = \frac{\sqrt{0.07A + \left(0.037L\sqrt{\dfrac{A}{n}}\right)}}{n}$$

(6.6)

where A = Area of study zone

L = Length of boundary of study zone

n = Number of individuals in study zone

These values are used in the z-test (eq. 6.4) in the same manner as for the Clark and Evans test. Donnelly (1978) suggests that this z-test is unbiased if $n > 7$ and the study area has a smooth boundary like a square or circle. It is not recommended if the study zone is a long, thin rectangle because the edge effects are overwhelming in this case.

Box 6.1 illustrates the use of the Donnelly modification of the Clark and Evans test for spatial pattern of tropical ant nests.

Box 6.1   Donnelly Modification of Clark and Evans Test for Spatial Pattern

Ant nests are typically thought by ecologists to be uniformly spaced in both tropical and temperate ant communities. Levings and Franks (1982) reported the following data from ground ants (*Ectatomma ruidum*) from one study area on Barro Colorado Island in Panama:

Study area = 100 m^2 $n = 30$ nests $L = 40$ m boundary length
Mean distance to nearest nest = 1.0863 m

They did not include a boundary strip around their plots, and we will use these data to illustrate the errors this introduces into the test of the null hypothesis of random spatial pattern of these ant nests.

1. Clark and Evans Test: from equation (6.2)

$\bar{r}_E$ = Expected distance to nearest neighbor if random pattern

$$= \frac{1}{2\sqrt{\rho}} = \frac{1}{2\sqrt{.30}} = 0.91287 \text{ meters}$$

The index of aggregation is thus (equation 6.3):

$$R = \frac{\bar{r}_A}{\bar{r}_E} = \frac{1.0863}{0.91287} = 1.19$$

which suggests a tendency toward a regular pattern of nest spacing. To test for a significant deviation from the expected random pattern we use equation (6.4):

$$z = \frac{\bar{r}_A - \bar{r}_E}{s_r} = \frac{1.0863 - 0.91287}{\left(0.26136 \middle/ \sqrt{(30)(0.30)} \right)} = 1.99$$

The value of z has probability 0.046 and consequently (at $\alpha = 0.05$) we reject the null hypothesis of a random spatial pattern for these ant nests in favor of a uniform pattern. This is an error because a boundary strip was not included in the study, and the correct statistical test is given next.

2. Donnelly modification of Clark and Evans Test: from equation (6.5)

$\bar{r}_c$ = Expected distance to nearest neighbor corrected for lack

of a boundary strip

$$= \bar{r}_E + \left[\left(0.051 + \frac{0.041}{\sqrt{n}} \right) \left(\frac{L}{n} \right) \right]$$

$$= 0.91287 + \left[\left(0.051 + \frac{0.041}{\sqrt{30}} \right) \left(\frac{40}{30} \right) \right] = 0.99085$$

The standard error of this expected distance is given by equation (6.6):

$$s_r = \frac{\sqrt{0.07A + \left(0.037 L \sqrt{\frac{A}{n}} \right)}}{n}$$

$$= \frac{\sqrt{0.07(100) + \left(0.037(40) \sqrt{\frac{100}{30}} \right)}}{30} = 0.10383$$

We use the same test as above (equation 6.4):

$$z = \frac{\bar{r}_A - \bar{r}_E}{s_r} = \frac{1.0863 - 0.99085}{0.10383} = 0.92$$

This value of z has probability 0.36 under the null hypothesis of a random pattern and we tentatively fail to reject the null hypothesis.

The Clarke and Evans test, if used without a boundary strip, biases the test for random patterns in the direction of uniform or regular patterns, as this example illustrates. In this case we have no evidence that supports the hypothesis that these ant nests are spaced in a uniform pattern, a pattern expected by most ant ecologists.

6.1.2 Distances to Second–n^{th} Nearest Neighbors

If distances to nearest neighbors are useful in revealing spatial patterns in

populations, additional information might be gained by measuring the distance to the

2nd, 3rd, 4th and n^{th} nearest neighbor. In practice it is difficult in the field to determine

easily who is the 3rd, 4th, and 5th nearest neighbor and no one seems to have gone beyond the 5th nearest neighbor. Thompson (1956) showed that the Clark and Evans approach could be easily generalized to the 2nd, 3rd... nearest neighbors. Table 6.1 gives the expected distances and their standard errors for the first to fifth nearest neighbors.

TABLE 6.1 Expected distances to first, second....fifth nearest neighbors and associated standard errors for a large population with a random pattern[a]

	Nearest neighbor				
Parameter	First	Second	Third	Fourth	Fifth
Expected mean distance	$\dfrac{0.5000}{\sqrt{\rho}}$ [b]	$\dfrac{0.7500}{\sqrt{\rho}}$	$\dfrac{0.9375}{\sqrt{\rho}}$	$\dfrac{1.0937}{\sqrt{\rho}}$	$\dfrac{1.2305}{\sqrt{\rho}}$
Standard error of expected distance	$\dfrac{0.2614}{\sqrt{n\rho}}$	$\dfrac{0.2723}{\sqrt{n\rho}}$	$\dfrac{0.2757}{\sqrt{n\rho}}$	$\dfrac{0.2774}{\sqrt{n\rho}}$	$\dfrac{0.2821}{\sqrt{n\rho}}$

[a] A boundary strip as in Figure 5.6 is assumed.

[b] ρ = Population density = Number of individuals/Area of study = n/A

Source: After Thompson, 1956.

Two additional tests are available for testing spatial pattern in nearest neighbor data.

Thompson's Test:

Thompson (1956) suggested a chi-square test for evaluating the null hypothesis of random spatial patterning. This test can be applied to nearest-neighbor data, as discussed above for the Clark and Evans test, or for more general studies of second nearest neighbor, third nearest neighbor and so on, and every one of these tests is independent. It is possible for nearest neighbors to be uniformly spaced while second nearest neighbors may be clumped in their distribution. Thompson's test is given by:

$$\chi^2 = 2\pi\rho \sum_{i=1}^{n}(r_i^2) \qquad (\text{d.f.} = 2nk)$$

(6.7)

where ρ = Population density on study area = n/A

r_i = Distance to k-th nearest neighbor for individual i

k = Rank of neighbors being measured (k = 1 for nearest neighbor,

       $k = 2$ for second nearest neighbor etc.)

     $n =$ number of individuals measured

     $\pi = 3.14159$

For example, for the cricket data given on page 239 above (with $\sum r_i^2 = 136.78$)

$$\chi^2 = 2(3.14159)(0.08621)(136.78)$$
$$= 74.09 \qquad [\text{d.f.} = 2(39)(1) = 78]$$

This value of χ^2 must be tested against two alternatives: a significantly *small* value of χ^2 indicates a *clumped* pattern and a significantly *large* value of χ^2 indicates a *uniform* pattern. For this example at $\alpha = .05$ we have two decision rules:

1. If observed χ^2 is less than $\chi^2_{.975}$ (59.0 in this example) we have evidence suggesting a clumped pattern.

2. If observed χ^2 is greater than $\chi^2_{.025}$ (99.6 in this example) we have evidence suggesting a uniform pattern.

In this example neither is true and for the cricket data we cannot reject the hypothesis of random patterning.

For large numbers of degrees of freedom, the usual chi-square text tables are not adequate and you need to use the normal approximation:

$$z = \sqrt{2\chi^2} - \sqrt{4nk - 1} \qquad\qquad (6.8)$$

where n and k are as defined above and z is the standard normal deviate. Negative values of z indicate a tendency to aggregation, positive values a tendency toward regularity of pattern. For this cricket example:

$$z = \sqrt{2(74.09)} - \sqrt{4(39)(1) - 1}$$
$$= -0.28$$

which is not significant ($p > .78$) and thus we tentatively fail to reject the null hypothesis of a random pattern.

If distances to the 2nd, 3rd ... nth nearest neighbor are available, Thompson's test is applied n times. It is possible that nearest-neighbors are uniformly distributed but that 5th nearest neighbors are clumped, so each test must be done independently.

Thompson's test is not as powerful as the following test if sample size is large ($n > 50$).

Goodness of fit Test:

Campbell and Clark (1971) proposed a chi-square goodness-of-fit test to compare the observed and expected distribution of nearest-neighbor distances. To use this test you must first group the observed measurements of nearest neighbor distances into a frequency distribution (see Box 6.2 for an illustration). The class limits for the frequency distribution must be chosen in advance, and this decision is arbitrary. Since the chi-square test requires an expected frequency in each cell of 3, the class limits picked should not be too narrow. A useful range would be to have about 5-10 classes in your frequency distribution, and this will require a large sample size of 50-100 measurements of nearest neighbor distances.

The expected frequency in each class of distances can be estimated by a two step procedure:

1. Estimate the cumulative probability for distances from 0 to r_i (where r_i is the upper class limit for each class):

$$F_x = 1 - e^{-x/2}$$
(6.9)

where

$$x = 2\pi \rho r_i^2$$
ρ = Population density
F_x = Probability of obtaining a distance in the range 0-r_i

For example, if the first class of nearest neighbor distances range from 0 to 0.85 m and the second class from 0.85 to 1.65 m, then for the cricket data used above where density is 0.08621:

For class 1: $x = 2(3.14159)(0.08621)(0.85^2) = 0.39136$

$$F_x = 1 - e^{-(0.39/2)} = 0.178$$

For class 2: $x = 2(3.14159)(0.08621)(1.65^2) = 1.4747$

$$F_x = 1 - e^{-(1.47/2)} = 0.522$$

Calculate F_x values for all the classes in the frequency distribution.

   2. Estimate the expected frequency in each class: These expected values can be obtained directly by subtraction as follows. Thus class 1 should contain 17.8% of the observed distances and class 2 should contain (52.2 - 17.8%) or 34.4% of the observed distances. With a sample size of 39 in this example, we should expect (39)(17.8%) or 6.94 distances to fall in the 0 – 0.85 m category and (39)(34.4%) or 13.42 distances to fall in the 0.85 – 1.65 m category. These expected frequencies should be calculated for all classes in the frequency distribution.

   The goodness-of-fit test follows directly from these calculations and is of the usual form:

$$\chi^2 = \sum \frac{(\text{Observed - expected})^2}{\text{Expected}} \qquad (\text{d.f.} = \text{no. of classes-1})$$

(6.10)

The only problem is to make sure to combine classes in the tails of the frequency distribution so that the expected value in each class is at least 3.

   Box 6.2 illustrates these calculations and Program PATTERN in Appendix 2 (page 000) can do these calculations.

Box 6.2 Campbell and Clarke Test of Goodness-of-fit for detecting spatial patterning in known populations

Nearest-neighbor distances were measured in a low-density cricket population in Australia by Campbell and Clarke (1971). They obtained these data on March 15, 1968 on a totally enumerated population: $n = 51$, $\sum r^2 = 4401.05$, $\bar{r}_A = 7.32$m.

Their distances ranged from 0.02 m to 23 m and they divided this range into 7 classes in intervals of 3.5 meters and tallied their raw data as follows

Distance to nearest neighbor-class limits (m)	Observed frequency
0-3.55	15
3.56-7.05	16
7.06-10.55	6
10.56-14.05	9
14.06-17.55	1

17.56-21.05	1
21.06- ∞	3

The study area was 17,800 m^2 and hence the known density was 51/17,800 or 0.002865 crickets/m^2.

1. Determine the expected frequency for these cricket data using equation (6.9) to fill in this table:

Class no.	Upper class limit (m)	X	$e^{-x/2}$	Cumulative probability $F_X = 1 - e^{-X/2}$
1	3.55	0.2269	0.8928	0.1072
2	7.05	0.8948	0.6393	0.3607
3	10.55	2.0037	0.3672	0.6328
4	14.05	3.5537	0.1692	0.8308
5	17.55	5.5448	0.0625	0.9375
6	21.05	7.9769	0.0185	0.9815
7	∞	∞	0.0	1.0000

The density of crickets (ρ) was 51 individuals on 17,800 m^2 or 2.865 X 10^{-3} individuals/m^2

For class 1:

$$x = 2\pi\rho r_1^2 = 2(3.14159)(0.002865)(3.55^2)$$
$$= 0.2269$$

and

$$F_x = 1 - e^{-X/2} = 1 - (2.718^{-0.1134})$$
$$= 0.1072$$

For class 2:

$$x = 2(3.14159)(0.002865)(7.05^2)$$
$$= 0.8948$$

and

$$F_x = 1 - (2.718^{-0.4474})$$
$$= 0.3607$$

similarly for classes 3-7, with the results shown in the above table.

2. We calculate the expected frequencies from the sample size ($n = 51$) and the formula

$$\left\{\begin{array}{l}\text{Expected relative} \\ \text{frequency in} \\ \text{class } x+1\end{array}\right\} = \left(\begin{array}{l}\text{Cumulative} \\ \text{probability in} \\ \text{class } x+1\end{array}\right) - \left(\begin{array}{l}\text{Cumulative} \\ \text{probability in} \\ \text{class } x\end{array}\right)$$

with the expected relative frequency in class 1 being equal to F_x for class 1.

Class no.	Cumulative probability F_x	Expected relative frequency	Expected no. of distances
1	0.1072	0.1072	5.47
2	0.3607	0.2535	12.93
3	0.6328	0.2721	13.88
4	0.8308	0.1980	10.10
5	0.9375	0.1067	5.44
6	0.9815	0.0440	2.24
7	1.000	0.0185	0.95
Total		1.0000	51.00

The expected relative frequencies must sum to 1.00 if you have done the calculations correctly.

To get the final expected number of distances, use the formula

$$\left\{\begin{array}{l}\text{Expected number} \\ \text{of distances}\end{array}\right\} = \left(\begin{array}{l}\text{Expected relative} \\ \text{frequency}\end{array}\right)\left(\begin{array}{l}\text{Sample} \\ \text{size}\end{array}\right)$$

For example, with $n = 51$:

$$\left\{\begin{array}{l}\text{Expected number of distances} \\ \text{in class 1}\end{array}\right\} = (0.1072)(51) = 5.47$$

$$\left\{\begin{array}{l}\text{Expected number of distances} \\ \text{in class 2}\end{array}\right\} = (0.2535)(51) = 12.93$$

etc. for the other classes.

We can now compute the chi-square test statistic:

$$\chi^2 = \sum \frac{(\text{Observed - expected})^2}{\text{Expected}}$$

taking care to combine the last two classes so that the expected value is above 3.

Class (m)	Observed no. of crickets	Expected no.	$\frac{(O - E)^2}{E}$
0-3.55	15	5.47	16.60
3.55-7.05	16	12.93	0.73
7.05-10.55	6	13.88	4.47
10.55-14.05	9	10.10	0.12
14.05-17.55	1	5.44	3.62
17.55-21.05	1	2.24	0.21
21.05-∞	3	0.95	
		Total	$\chi^2 = 25.75$

with (6 - 1) degrees of freedom. Since the critical value of χ^2 at $\alpha = 0.01$ is 15.09 for 5 d.f., this value is highly significant and we reject the null hypothesis of random spacing. Since there are more crickets at short nearest-neighbor distances than expected, the data suggest a clumped distribution.

The nearest neighbor subprogram in Program PATTERN (Appendix 2, page 000) will do these calculations from the raw data of nearest-neighbor distances.

6.1.3 More Sophisticated Techniques for Spatial Maps

If an ecologist has a spatial map of a population (as in Fig. 6.2) it is possible to achieve a much more detailed description of the map than one obtains with nearest-neighbor data. In most cases special stochastic models are fitted to map data, and such models are at a level of mathematical complexity that is beyond the scope of this book. I sketch here only the outlines of this approach and students are referred to Ripley (1981), Diggle (1983), and Perry (1995) for further details.

A spatial map of the locations of animals or plants can be represented by points on a plane (Fig. 6.3). We can assign to each individual organism a "territory" which includes all the space which is closer to this individual than it is to any other individual. These "territories" should not be confused with the home ranges or territories of vertebrates; they are perhaps better thought of in terms of plants competing for soil

water or soil nutrients. This mathematical construction of "territories" is called the *Dirichlet tessellation* of the organisms in the study area, and is illustrated in Figure 6.3.
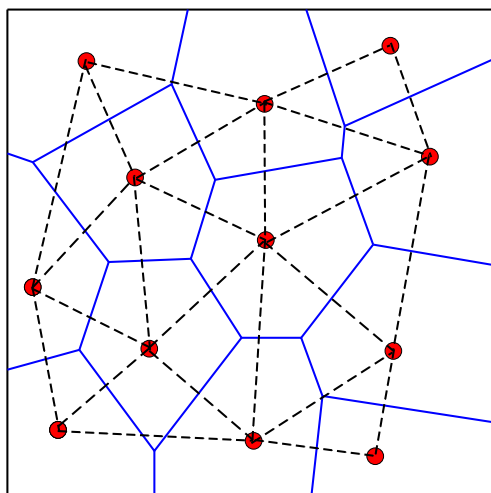

Figure 6.3 Hypothetical spatial map of a population of 12 organisms (red dots) in a study area. The *Dirichlet tessellation* (blue lines) assigns a "territory" to each organism that contains all the space that is closer to that organism than it is to any other. The resulting polygons are shown in blue. Lines that join all individuals with a common boundary to their territory define a set of triangles called the *Delaunay triangulation* (dashed lines). The sides of these triangles are a set of nearest-neighbor distances.

Individuals that share a common boundary of their "territories" can be connected with lines that are the equivalent of the nearest-neighbor distances discussed above. These lines connect to form a series of triangles called the *Delaunay triangulation*, as shown in Figure 6.3.

The computation of the Dirichlet tessellation becomes much more complicated as the number of individuals in the study area increases. We shall not attempt these calculations in this book, and interested students should refer to Diggle (1983) and Perry (1995).

With a spatial map like Figure 6.3 we can compute *all* of the distances between each individual and every other individual in the study area. Using Monte Carlo simulations with a computer, we can compare the observed distances to those expected under complete spatial randomness (see Diggle 1983, pp. 11-16). We can also locate random points on the map and measure point-to-nearest-organism distances and compare these with computer-generated expected values. There is an

extensive literature on the analysis of mapped patterns and the fitting of stochastic models to such data. Consult Pielou (1977), Diggle (1983), and Ripley (1981) for the details of this approach which moves fairly rapidly into difficult mathematics.

The analysis of spatial maps becomes even more complex when the study area is not spatially homogenous. For example, birds nesting in isolated patches of forest often position their nests near the edge of the wood. If small blocks of woodland are patchily distributed, the nearest neighbor of a nest in one wood might be another nest in the next patch of forest. Edge effects in patchy environments can cause severe biases in nearest-neighbor methods, and the common sense notion of discarding all nearest-neighbor distances that cross habitat boundaries is not a solution (Ripley 1985). One possible method of approach useful in patchy environments is outlined in Ripley (1985), but the problem awaits more theoretical analysis.

The advent of Geographic Information Systems (GIS) has produced a whole new field of analysis of map data, and has evolved into a speciality of its own. An introduction to some methods are given in Scheldeman and van Zonneveld (2010), which is available on the web at (http://www.bioversityinternational.org/index.php?id=19&user_bioversitypublications_pi1=5052) . Consulting a GIS specialist is recommended if you have spatial map data for analysis.

6.2 CONTIGUOUS QUADRATS

Quadrats are widely used in plant ecology to analyze the spatial patterning of vegetation, as we have already seen in Chapter 4. Here we introduce a new type of quadrat sampling based on adjacent or contiguous quadrats. In this case typically a series of contiguous quadrats are counted as illustrated in Figure 6.4. These quadrats are clearly not independent random samples from the population, and consequently they cannot be used to estimate population density or percent cover or any other measure of abundance. They are best viewed as a statistical universe about which questions of pattern may be asked and the major operational question is where to start the quadrats.
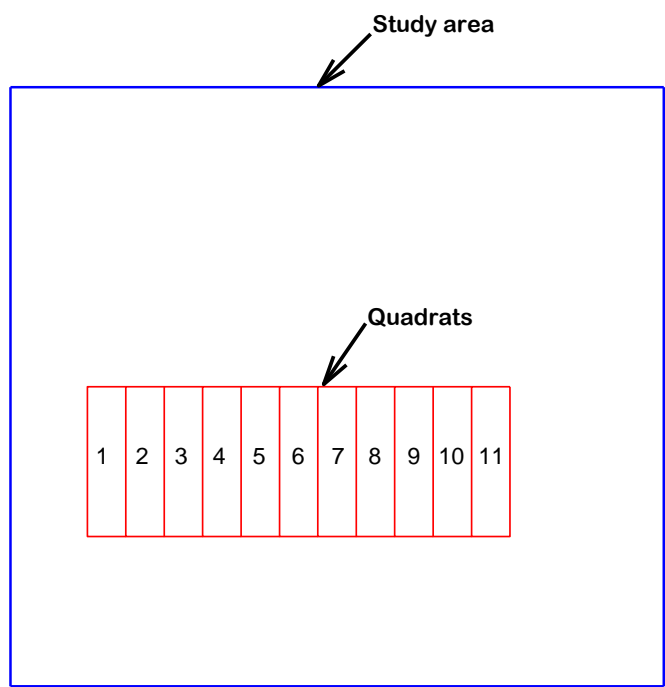
Figure 6.4 Schematic illustration of a series of 11 contiguous quadrats (red) for the analysis of statistical pattern in plant communities. Such quadrats are not independent and cannot be used to estimate abundance or cover, but they can be used to determine spatial pattern.

There are two main methods of analyzing data from contiguous quadrats, and these are discussed in detail by Fortin and Dale (2005). I will present here only the simplest method, the blocked quadrat variance method developed by Hill (1973). The specific methods developed by Hill (1973) have been given the acronym TTLQV (two-term local quadrat variance), and are applied to data gathered from a series of contiguous quadrats like those shown in Figure 6.4. Blocks are defined by adding together adjacent quadrat data as follows:

Block size	Quadrats
1	(1) (2) (3) (4) (5) (6) (7) (8)
2	(1,2) (3,4) (5,6) (7,8)
3	(1,2,3) (4,5,6)
4	(1,2,3,4) (5,6,7,8)

Given these grouping of data into block sizes we proceed to calculate the variances for each block size, as follows:

$$\text{Var}_1(X) = \left(\frac{1}{n-1}\right)\left[\left(\frac{(x_1 - x_2)^2}{2}\right) + \left(\frac{(x_2 - x_3)^2}{2}\right) + \cdots + \left(\frac{(x_{n-1} - x_n)^2}{2}\right)\right] \qquad (6.11)$$

$$\text{Var}_2(X) = \left(\frac{1}{n-3}\right)\left[\begin{array}{c}\left(\dfrac{(x_1 + x_2 - x_3 - x_4)^2}{4}\right) + \left(\dfrac{(x_2 + x_3 - x_4 - x_5)^2}{4}\right) + \cdots \\[4mm] \cdots + \left(\dfrac{(x_{n-3} + x_{n-2} - x_{n-1} - x_n)^2}{4}\right)\end{array}\right] \qquad (6.12)$$

where $\text{Var}_1(X)$ = Variance of counts at block size 1

$\text{Var}_2(X)$ = Variance of counts at block size 2

n = Number of quadrats

x_1 = Number of organisms counted in quadrat 1, etc.

and these calculations are carried forward in a similar manner for block sizes 3, 4, 5 ... to the upper limits of the possible pooling of quadrats. The upper limit of block size is $n/2$ but the recommended upper limit is $n/10$. For example, if you counted 150 contiguous quadrats, you could calculate block sizes for up to two blocks of 75 quadrats each, but it is recommended that you stop at block size 15. A plot of these variances against the block size can be used to determine the spatial pattern shown in the species being analyzed.

6.2.1 Testing for Spatial Pattern

After computing the variances from grouped sets of adjacent quadrats (TTLQV method), we can now ask how to use these estimates to determine spatial pattern. A plot of the variances (eq. 6.11) against block size for the TTLQV method will show in general one of three patterns corresponding to the spatial patterns of the organisms (Figure 6.5). If the individuals are dispersed at random over the study zone, the plot of variance against block size will fluctuate irregularly with no pattern. If the spacing pattern is uniform, the variances estimated will all be low and will not tend to fluctuate with block size. Finally, if the individuals are clumped in their distribution, the variances will tend to peak at a block size equivalent to the radius of clump size (the

average area occupied by a clump). The average distance between the clumps will be twice this block size. If the variance peak is high and sharp, the clumps are tight and the clumped pattern is said to be of high intensity with distinct clumps and large open spaces between the clumps. If the variance peak is low, the pattern is of low intensity and the clumps are not well defined. Since many organisms are clumped in natural systems, the TTLQV approach can be used effectively to define and quantify the type of clumping. Box 6.3 provides an illustration of this approach.
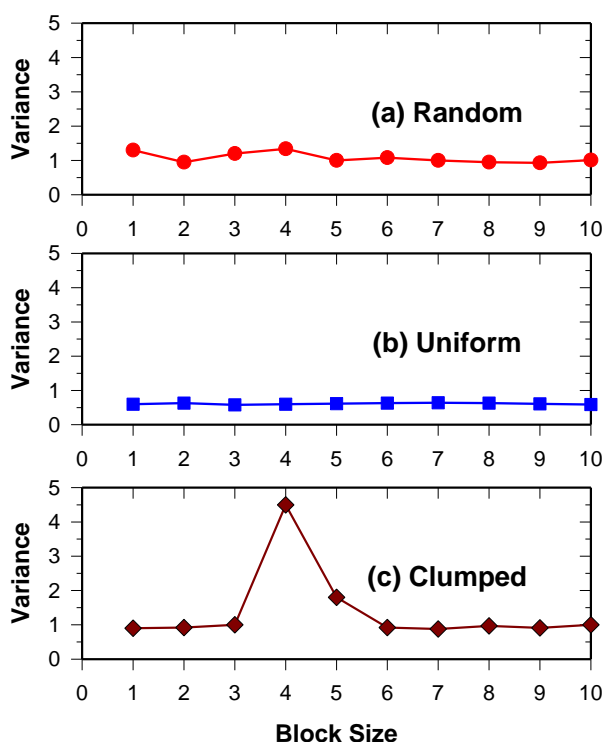
Figure 6.5 Schematic illustration of the type of plots that are found for Hill's TTLQV method for contiguous quadrats. A plot of the block variances against the size of the blocks will give different patterns for (a) random, (b) uniform, and (c) clumped distributions.

Box 6.3   Calculation of variances for continuous quadrats using Hill's TTLQV method

Amabilis fir seedlings were counted in a series of 104 contiguous quadrats, each 10 by 10 cm. The seedling counts obtained were as follows:

```
0 0 1 0 5 7 2 1 0 0 0 1 0 4 6 3 0 0 1 1 0 0 0 2 5 8
0 2 4 6 5 2 1 0 0 1 2 4 7 3 0 0 1 0 1 0 3 6 5 2 0 0
0 1 3 8 4 1 0 0 1 1 0 5 6 3 1 0 0 0 1 0 1 4 7 4 2 0
1 2 0 3 4 6 4 0 1 0 1 0 3 6 7 5 2 0 1 0 0 0 2 3 7 4
```

1. Calculate the variance for block size 1 from equation (6.11):

$$\text{Var}_1(X) = \left(\frac{1}{n-1}\right)\left[\left(\frac{(x_1 - x_2)^2}{2}\right) + \left(\frac{(x_2 - x_3)^2}{2}\right) + \cdots + \left(\frac{(x_{n-1} - x_n)^2}{2}\right)\right]$$

$$\text{Var}_1(X) = \left(\frac{1}{104-1}\right)\left[\left(\frac{(0-0)^2}{2}\right) + \left(\frac{(0-1)^2}{2}\right) + \cdots + \left(\frac{(7-4)^2}{2}\right)\right]$$

$$= 2.5922$$

2. Calculate the variance for block size 2 from equation (6.12):

$$\text{Var}_2(X) = \left(\frac{1}{n-3}\right)\left[\left(\frac{(x_1 + x_2 - x_3 - x_4)^2}{4}\right) + \left(\frac{(x_2 + x_3 - x_4 - x_5)^2}{4}\right) + \cdots + \left(\frac{(x_{n-3} + x_{n-2} - x_{n-1} - x_n)^2}{4}\right)\right]$$

$$\text{Var}_2(X) = \left(\frac{1}{104-3}\right)\left[\left(\frac{(0+0-1-0)^2}{4}\right) + \left(\frac{(0+1-0-5)^2}{4}\right) + \cdots + \left(\frac{(2+3-7-4)^2}{4}\right)\right]$$

$$= 7.9901$$

3. Continue these calculations to block size 10 (approximate n/10) with the following results:

$\text{Var}_3(X) = 12.3305$

$\text{Var}_4(X) = 13.7590$

$\text{Var}_5(X) = 11.8579$

$\text{Var}_6(X) = 8.2294$

$\text{Var}_7(X) = 5.2166$

$\text{Var}_8(X) = 3.4228$

$\text{Var}_9(X) = 2.8186$

$\text{Var}_{10}(X) = 2.7641$

4. These variances can be plotted as a function of block size, as in Figure 6.5c. The resulting pattern shows a strong peak of variance in the range of block sizes 3 to 5. The ecological interpretation is that Amabilis fir seedlings in this community are strongly clumped, and on average the clumps are 60-100 cm apart (twice the block size peak multiplied by the quadrat size of 10 cm).

Program PATTERN (Appendix 2, page 000) can do these calculations.

6.3 SPATIAL PATTERN FROM DISTANCE METHODS

In many cases the entire population can not be mapped as it is in Figure 6.2 and we must sample individuals scattered over a large area. We have discussed the use of distance methods to estimate population density. We now wish to use these same methods to determine spatial pattern. Using the same general approach shown in Figure 5.6, we can make two kinds of measurements:

> **1.** from random *points* to the nearest organism
> **2.** from a random *organism* to its nearest neighbor.

6.3.1 T-Square Sampling Procedure

Given the T-square distance data illustrated in Figure 5.7, we can test the hypothesis of a random spatial pattern in the population. The most powerful test statistic recommended by Hines and Hines (1979) is:

$$h_T = \frac{2n[2 \sum (x_i^2) + \sum (z_i^2)]}{[(\sqrt{2} \sum x_i) + \sum z_i]^2}$$

(6.13)

where h_T = Hines' test statistic for randomness of T-square data
n = Sample size (no. of random points)
x_i = Point-to-organism distances
z_i = T-square organism-to-neighbor distances

This test statistic is evaluated by referring to critical values in Table 6.2. Low values of h_T indicate a uniform pattern, and high values indicate aggregation. In a random world h_T is 1.27, and smaller values indicate a uniform pattern while larger values indicate clumping.

Diggle (1983) and Hines and Hines (1979) consider other test statistics for T-square data, but none of them seems superior to h_T in statistical power.

6.3.2 Eberhardt's Test:

One test for spatial pattern and associated index of dispersion that can be used on random-point-to-nearest-organism distances was suggested by Eberhardt (1967) and analyzed further by Hines and Hines (1979):

$$I_E = \left(\frac{s}{\bar{x}}\right)^2 + 1$$

(6.14)

where I_E = Eberhardt's index of dispersion for point-to-organism distances
s = Observed standard deviation of distances
$\bar{x}$ = Mean of point-to-organism distances

This index does not depend on the density of the population, and it is easy to calculate. For example, using the spruce tree data in Box 5.2 (page 218):

$$n = 16; \quad \bar{x} = 9.681; \quad s^2 = 13.020$$

we get:

$$I_E = \left(\frac{\sqrt{13.020}}{9.681} \right)^2 + 1 = 1.14$$

Critical values of I_E have been computed by Hines and Hines (1979) and are given in Table 6.2 (page 000), which is entered with sample size $2n$. The expected value of I_E in a random population is 1.27. Values below this suggest a regular pattern, and larger values indicate clumping.

TABLE 6.2 Critical values for the Hines test statistic h_T (eq. 6.13) which test the null hypothesis that spatial pattern is random in a population sampled with the T-square sampling procedure illustrated in Figure 5.7[a]

	Regular alternative				Aggregated alternative			
$n/\square$	0.005	0.01	0.025	0.05	0.05	0.025	0.01	0.005
5	1.0340	1.0488	1.0719	1.0932	1.4593	1.5211	1.6054	1.6727
6	1.0501	1.0644	1.0865	1.1069	1.4472	1.5025	1.5769	1.6354
7	1.0632	1.0769	1.0983	1.1178	1.4368	1.4872	1.5540	1.6060
8	1.0740	1.0873	1.1080	1.1268	1.4280	1.4743	1.4743	1.5821
9	1.0832	1.0962	1.1162	1.1344	1.4203	1.4633	1.4539	1.5623
10	1.0912	1.1038	1.1232	1.1409	1.4136	1.4539	1.4456	1.5456
11	1.0982	1.1105	1.1293	1.1465	1.4078	1.4456	1.4384	1.5313
12	1.1044	1.1164	1.1348	1.1515	1.4025	1.4384	1.4319	1.5189
13	1.1099	1.1216	1.1396	1.1559	1.3978	1.4319	1.4261	1.5080
14	1.1149	1.1264	1.1439	1.1598	1.3936	1.4261	1.4209	1.4983
15	1.1195	1.1307	1.1479	1.1634	1.3898	1.4209	1.4098	1.4897
17	1.1292	1.1399	1.1563	1.1710	1.3815	1.4098	1.4008	1.4715
20	1.1372	1.1475	1.1631	1.1772	1.3748	1.4008	1.3870	1.4571

25	1.1498	1.1593	1.1738	1.1868	1.3644	1.3870	1.3768	1.4354
30	1.1593	1.1682	1.1818	1.1940	1.3565	1.3768	1.3689	1.4197
35	1.1668	1.1753	1.1882	1.1996	1.3504	1.3689	1.3625	1.4077
40	1.1730	1.1811	1.1933	1.2042	1.3455	1.3625	1.3572	1.3981
45	1.1782	1.1859	1.1976	1.2080	1.3414	1.3572	1.3528	1.3903
50	1.1826	1.1900	1.2013	1.2112	1.3379	1.3528	1.3377	1.3837
75	1.1979	1.2043	1.2139	1.2223	1.3260	1.3377	1.3260	1.3619
100	1.2073	1.2130	1.2215	1.2290	1.3189	1.3289	1.3189	1.3492
150	1.2187	1.2235	1.2307	1.2369	1.3105	1.3184	1.3105	1.3344
200	1.2257	1.2299	1.2362	1.2417	1.3055	1.3122	1.3055	1.3258
300	1.2341	1.2376	1.2429	1.2474	1.2995	1.3049	1.2995	1.3158
400	1.2391	1.2422	1.2468	1.2509	1.2960	1.3006	1.2960	1.3099
500	1.2426	1.2454	1.2496	1.2532	1.2936	1.2977	1.2936	1.3059

[a] n = Sample size. Low values of h_T indicate a clumped pattern.

Source: Hines and Hines, 1979

Eberhardt's index can be applied to point-quarter data by using only the nearest of the four distances measured at each point in equation (5.34). This is very inefficient compared with the previous methods.

6.3.3 *Variable Area Transects*

The variable area transect method described in Chapter 5 can also provide data for a test of the null hypothesis of a random spatial pattern. Parker (1979) suggests a statistical test by calculating for each observation two values:

$$P_i = \frac{i}{n}$$

$$S_k = \frac{\sum\limits_{i=1}^{k} l_i}{\sum\limits_{i=1}^{n} l_i}$$

(6.15)

where P_i = Proportion of sample counted up to quadrat I

         n = Total numbers of quadrats sampled

         S_k = Proportion of total lengths measured up to quadrat k

         l_i = Length to third organism for quadrat number I (I = 1, 2, 3, 4, ….n)

Under the null hypothesis of randomness, the two proportions P_i and S_k will rise together. To test the null hypothesis search for the largest deviation between the two proportions, and refer this maximum deviation to the table for the Kolmolgorov-Smirnov goodness of fit test (Zar 1996, page 474 and Table B9; Sokal and Rohlf 1995, page 708).

Figure 6.6 illustrates this test for randomness, and Box 6.4 provides an example of the calculations.
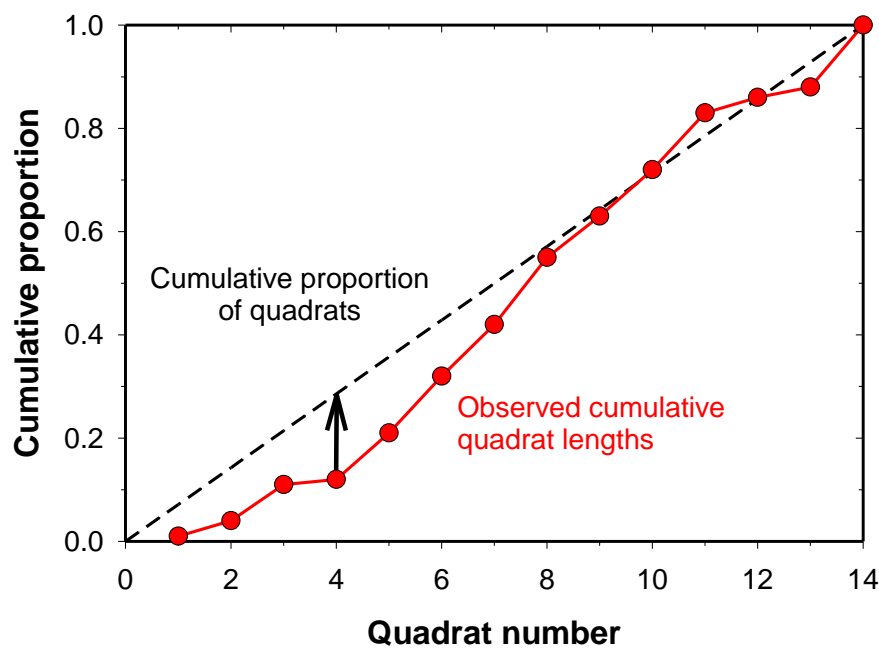

Figure 6.6 Test of the null hypothesis of random spatial pattern for the variable area transect method of Parker (1979). Two cumulative distributions are plotted for the sample data in Box 6.4 on pismo clams (*Tivela stultorum*) in California. The dashed line gives the cumulative proportion of quadrats searched and is a straight line. The data points are the cumulative proportion of the total quadrat lengths searched up to each quadrat number. The maximum deviation between these two distributions (arrow) is the Kolmogorov-Smirnov statistic *D* tabled in Zar (1996, Table B.9).

Box 6.4 Parker's test for random spatial pattern with variable area transect data

Parker (1979) reported the following data on the abundance of pismo clams (*Tivela stultorum*) at Seal Beach, California. The transect width was 0.23 m and the following data was obtained, following the procedures illustrated in Figure 5.9:

Quadrat no.	Length of quadrat searched to reach the third clam, l_i (m)	Cumulative proportion of quadrats searched, P_i	Cumulative proportion of lengths measured, S_k
1	5.4	0.071	0.011
2	15.3	0.143	0.041
3	33.2	0.214	0.106
4	7.8	0.286	0.121
5	43.0	0.357	0.205
6	60.5	0.429	0.323
7	49.8	0.500	0.421
8	63.4	0.571	0.545
9	42.3	0.643	0.628
10	35.8	0.714	0.698
11	65.8	0.786	0.827
12	15.3	0.857	0.857
13	13.9	0.929	0.884
14	59.3	1.000	1.000

1. To calculate the cumulative proportion of quadrats searched, use equation (6.15):

$$P_i = \frac{i}{n}$$

$$P_1 = \frac{1}{14} = 0.071 \qquad P_2 = \frac{2}{14} = 0.143 \qquad \cdots\cdots$$

2. To calculate the cumulative proportion of lengths measured, we need to sum the 14 lengths:

$$\sum l_i = 5.4 + 15.3 + 33.2 + \cdots\cdots = 510.8 \text{ meters}$$

From this, using equation (6.15) we obtain:

$$S_k = \frac{\sum_{i=1}^{k} l_i}{\sum_{i=1}^{n} l_i}$$

$$S_1 = \frac{5.4}{510.8} = 0.011$$

$$S_2 = \frac{5.4 + 15.3}{510.8} = 0.041 \quad \cdots\cdots$$

These calculations fill in the third and fourth columns of the above table.

3. Calculate the difference between these two cumulative frequency distributions:

$$d_i = |P_i - S_i|$$

where d_i = Differences between the two frequency distributions for quadrat i

For these data it is clear the the largest difference between the cumulative proportions is at sample 4 where $d_4 = |0.286 - 0.121| = 0.165$

From Table B.9 in Zar (1996) the critical value of the Kolmogorov-Smirnov test statistic at $\alpha = 0.05$ is 0.349 for $n = 14$, and since our observed value is less than the tabled value, we tentatively accept the null hypothesis of a random pattern.

In practice of course this sample size is very small, and you would be advised to try to measure 30-40 variable area quadrats to have a more powerful test.

6.4 INDICES OF DISPERSION FOR QUADRAT COUNTS

Many populations of animals and plants are aggregated in nature, and a few are spaced out in a regular pattern. One reason for knowing these patterns is that they affect decisions about what method to use for estimating population density, as we have seen. But a second reason for knowing the pattern is to describe it objectively and to try to explain it biologically. In this section we will describe various measures that have been proposed for quadrat counts to quantify pattern in natural populations.

A series of counts of the numbers of individual animals or plants are taken in n quadrats of defined size and shape, as discussed in Chapter 4. We wish to use these counts to construct an index of dispersion, all the while remembering that an index of dispersion will depend on quadrat size and shape. All indices have their defects but some are still better than others. I will discuss five possible indices of dispersion here. Elliott (1977) and Myers (1978) discuss several other indices but none seems very useful in practice.

6.4.1 Variance-to-mean ratio

This is one of the oldest and one of the simplest measures of dispersion. The ratio ($s^2 / \bar{x}$) is usually called the *index of dispersion* (*I*) (equation 4.3, page 151) and is based on the observation that in a random pattern, described by the Poisson distribution, the variance equals the mean, so $I = 1$. There are clear problems with the index of dispersion as a measure of dispersion.

The use of the index of dispersion as a measure of spatial pattern was criticized by Hurlbert (1990) and by Perry (1995). The essence of these criticisms is that there are possible non-random patterns that produce variance-to-mean ratios of 1. Two examples from quadrat counts are:

(a)

(b)

2	6	6	$n = 9$
2	6	6	$\bar{x} = 4$
2	2	4	$s^2 = 4$

8	6	3	$n = 9$
5	4	2	$\bar{x} = 4$
3	3	2	$s^2 = 4$

Variance/mean ratio = 1.0   Variance/mean ratio = 1.0

Clearly both these patterns are non-random, (a) being aggregated and bimodal and (b) showing a strong gradient. This non-randomness is not captured by the variance-to-mean ratio. Other measures of dispersion need to be considered and if your interest is in measuring spatial pattern the recommendation is not to use the index of dispersion (although many papers in the literature make this mistake).

6.4.2 k of the negative binomial:

The exponent k is frequently used as an index of dispersion for sets of data that fit a clumped pattern and are adequately described by the negative binomial distribution (see Chapter 4, pg. 000). Table 6.3 gives the expected values for negative binomial k when there is a uniform or a clumped pattern The first requirement for the use of this index is that you must check your data for agreement with the negative binomial. Since small values of k indicate maximum clumping, it is customary to use the reciprocal ($1/k$) as the index of pattern, rather than k itself. Table 6.3 shows that $1/k$ is possibly affected by sample size. You should not use ($1/k$) as an index of pattern unless all the data sets that are being compared have the same sample size. Since this condition are may not often be correct, this index of dispersion is not typically recommended.

TABLE 6.3   Limits of three indices of spatial patterning in populations sampled by quadrat counts[a]

	Value expected under		
	Maximum uniformity	Randomness	Maximum aggregation
Reciprocal of k (negative binomial exponent)	$-\dfrac{1}{\bar{x}}$	0	$n - \dfrac{1}{\bar{x}}$
Green's coefficient[b]	$-\dfrac{1}{\sum(x) - 1}$	0	1
Morisita coefficient	$1 - \left(\dfrac{n-1}{\sum(X) - 1}\right)$	0	n

[a] Values of these indices are given for conditions of maximum uniformity, randomness, and maximum aggregation.

[b] Green's corefficient should only be used for aggregated populations.

n = sample size, X = observed quadrat counts, $\bar{x}$ = mean population density.

6.4.3 Green's Coefficient:

Green (1966) devise a coefficient of dispersion that would measure aggregation and be useful for data in which the variance/ mean ratio was greater than 1. His coefficient is simple to compute:

$$\text{Green's coefficient of dispersion} = \frac{(s^2 / \bar{x}) - 1}{\sum (X) - 1}$$

(6.16)

Green's coefficient should not be used for uniform patterns and positive values indicate a clumped pattern. Unfortunately the sampling distribution of Green's coefficient has not been worked out and it is difficult to assign confidence limits to it.

6.4.4 Morisita's Index of Dispersion:

Morisita (1962) developed an index of dispersion that has some of the desirable attributes of an index:

$$I_d = n \left[\frac{\sum x^2 - \sum x}{(\sum x)^2 - \sum x} \right]$$

(6.17)

where I_d = Morisita's index of dispersion

   n = Sample size

$$\sum x = \text{Sum of the quadrat counts} = x_1 + x_2 + x_3 \cdots$$

$$\sum x^2 = \text{sum of quadrat counts squared} = x_1^2 + x_2^2 + x_2^3 \cdots$$

This coefficient has the desirable statistical property of having a known sampling distribution. Morisita (1962) showed that one could test the null hypothesis of randomness by:

$$\chi^2 = I_d(\sum x - 1) + n - \sum x \qquad (\text{d.f.}= n{-}1) \tag{6.18}$$

where χ^2 = Test statistic for Morisita's index of dispersion (chi-square

distribution)

and other terms are defined above.

The Morisita Index is nearly identical to the index of mean crowding devised by Lloyd (1967) and for all practical purposes these two indices can be considered identical. Lloyd (1967) suggested that with quadrat counts one could estimate a parameter that was analogous to the amount of crowding an average individual would experience. He called this 'mean crowding' and defined it as for each individual plant or animal the mean number of other individuals within a sample quadrat. Mean crowding is most readily defined for quadrat samples that fit a negative binomial distribution. The sample estimates given by Lloyd (1967) are:

$$\overset{*}{X} = \bar{X} + \frac{\bar{X}}{k} \tag{6.19}$$

where $\overset{*}{X}$ = estimate of mean crowding

$\bar{X}$ = mean density per quadrat

k = negative binomial k (Chapter 4, page 159 ff.)

The large sample standard error of mean crowding can be estimated by:

$$\text{S.E.}\left[\overset{*}{X}\right] = \frac{\bar{X}}{k^2}\sqrt{\left[\text{var}(\hat{k}) + \frac{\hat{k}(\bar{X}+\hat{k})(1+\hat{k})^2}{q\bar{X}}\right]} \tag{6.20}$$

where $\text{var}(\hat{k})$ = sampling variance for negative binomial k

q = number of quadrats counted

Lloyd (1967) used the index of mean crowding to define a biologically more intuitive measure he called "patchiness", the ratio of mean crowding to mean density:

$$\frac{\overset{*}{X}}{\overline{X}} = \text{Lloyd's index of patchiness} = 1 + \frac{1}{\hat{k}} \tag{6.21}$$

which he defined as "how many times as 'crowded' an individual is, on the average, as it would have to be if the same population had a random distribution". Lloyd (1967) was concerned with estimates of spatial dispersion derived from quadrat counts that are aggregated and fit a negative binomial distribution. Lloyd's analysis points out the critical fact relevant to all measures of dispersion that they depend on quadrat size.

Hurlbert (1990) suggests that the Morisita Index (or its analogue Lloyd's index of patchiness) is one of the better indices of dispersion.

The sample size necessary to obtain a reliable index of dispersion is difficult to estimate because estimates of confidence limits are not available for most coefficients. Green (1966) recommends that a minimum sample size should be 50 quadrats, and that when the pattern is highly clumped, at least 200 quadrats are required.

6.4.4 Distance to Regularity Indices

Perry and Hewitt (1991) criticized the conventional approaches to measuring aggregation and uniformity because all the existing methods ignore the movements of organisms. Perry (1995a,1995b) developed two indices of dispersion based on the concept of distance-to-regularity. Given a set of quadrat counts, how many individuals would have to moved about to produce a set of counts that show maximum regularity. In a simple world maximum regularity would be shown by equal counts in all quadrats. This approach computes the number of organisms that have to be moved to produce regularity and compares it with the expected distance to regularity calculated by randomly and independently assigning individuals to quadrats in a computer with an equal probability of being assigned to each quadrat. The computer reclassification is typically done 1000 times to generate the expected distance to regularity. Perry (1995a) proposed the following index:

$$I_r = \frac{D}{E_r}$$

(6.22)

where I_r = index of regularity

D = observed distance to regularity

E_r = expected distance to regularity for the permuted data

This index is 1.0 when the sample is random, less than 1 when uniform, and greater than 1 when aggregated. This index can be calculated for any set of quadrats for which the spatial framework is known. Box 6.5 gives an illustration of this method.

Box 6.5 Calculation of Perry's Index of Regularity and Index of Aggregation

---

Monte Lloyd counted centipedes (*Lithobius crassipes*) on 6 quadrats equally spaced at 2 ft intervals and obtained these data:

2 1 2 4 5 10 with a total of 24 individuals in 6 quadrats

Quadrat A	B	C	D	E	F
2	1	2	4	5	10

We can compute two indices from these data.

1. Calculate the observed distance to regularity for the data.

In this simple example, a regular or uniform distribution would be:

4 4 4 4 4 4

To achieve this uniform pattern, we can hypothetically move centipede individuals one at a time from the original observed pattern. In this simple case:

move 2 individuals from quadrat F to quadrat A: 2×10 ft $= 20$ ft
move 3 individuals from quadrat F to quadrat B: 3×8 ft $= 24$ ft
move 1 individuals from quadrat F to quadrat C: 1×6 ft $= 6$ ft
move 1 individual from quadrat E to quadrat C: 1×4 ft $= 4$ ft

All quadrats in this hypothetical universe now have 4 centipedes. The total distance we had to move centipedes to achieve a regular patterns is:

$D = 20 + 24 + 6 + 4 = 54$ ft

2. We can reassign the 24 individuals to the 6 quadrats at random, with an equal probability of being put in any one of the 6 quadrats. For each random rearrangement we can compute the moves we would have to make to achieve a regular pattern. For example, one random toss for the individuals might be:

Quadrat A	B	C	D	E	F
2	3	7	3	4	5

By moving these hypothetical individuals in the direction of regularity, you can verify that for this arrangement you need to move:

$$4+2+4=10 \text{ ft}$$

By repeating this 400 times you can determine that the expected value of the distance moved to obtain a regular pattern is 21.9 ft. From equation (6.22):

$$I_r = \frac{D}{E_r} = \frac{54}{21.9} = 2.47$$

From the 400 random selections, only 0.7 % had values above the observed 54 ft, and this suggests a significantly aggregated pattern in these centipedes.

This index of regularity is 1.0 for random counts and above 1 for aggregated patterns. Note that this index does not take into account the position of the quadrats.

2. Perry (1995) developed a second index to measure aggregation when the position of the quadrat is important. For example there may be a gradient in the abundances. To test for these types of non-random patterns proceed as follows.

To calculate the expected distance to regularity, we need to randomly assign these counts to randomly selected quadrats. For example, one of the random rearrangements might be:

Quadrat A	B	C	D	E	F
4	2	10	1	2	5

Note that in this case we do not rearrange the individuals but rearrange the counts for the quadrats. This is equivalent to moving the quadrats in space and thus directly addresses the question of whether or not the geographic positions of the quadrat are effectively random. By moving these hypothetical individuals in the same way shown above, you can verify that for this arrangement you need to move:

$$4+6+4+2=16 \text{ ft}$$

If you repeat these permutations many times and average them you will find the expected distance to regularity is 29.3 ft. From equation (6.23):

$$I_a = \frac{D}{E_a} = \frac{54}{29.3} = 1.84$$

From 400 simulations on the computer only 1.5% of the values exceeded the observed 54 ft, and we can thus conclude that the spatial pattern of these centipedes is significantly aggregated. An index of aggregation of 1 indicates random patterns, and > 1 clumped patterns.

This is a simple case for illustration. A much larger sample size of 30-50 quadrats would be desirable in a field study of spatial pattern.

These calculations are clearly most easily done in a computer program, and Perry (1995) has developed Program SADIE to do these permutations.

A second index proposed by Perry (1995a) is based on a similar approach of distance to regularity but now instead of reassigning the individuals at random to

quadrats, you reassign the counts to different quadrats. This approach tests the spatial arrangement of the counts, and for example if there is a gradient among the quadrats (as shown above, page 262), it will be detected. The second index is:

$$I_a = \frac{D}{E_a}$$

(6.23)

where I_a = index of aggregation

D = observed distance to regularity

E_a = average distance to regularity for the permuted counts

Both these indices should be utilized for quadrat count data since they detect different forms of aggregation. Box 6.5 illustrates this approach.

Perry (1995b) has developed another method called SADIE (Spatial Analysis by Distance Indices) for the analysis of mapped data in which the exact coordinates of individuals are known. Again the principle of moving individuals is utilized, but now since the spatial map is known, it is possible to move them directionally and stepwise toward complete regularity (Figure 6.7). The observed number and direction of moves to regularity in the data are then compared to the simulated number and direction of moves needed to change a randomly generated map into a regular pattern. An index of pattern identical to those given above can be generated with this approach. For example, the redwood seedling map in Figure 6.2 provides an index of pattern of 1.28, and shows significant aggregation. The mathematical details are given in Perry (1995b), and the procedures are performed by Program SADIEM as described in Perry (1995b).
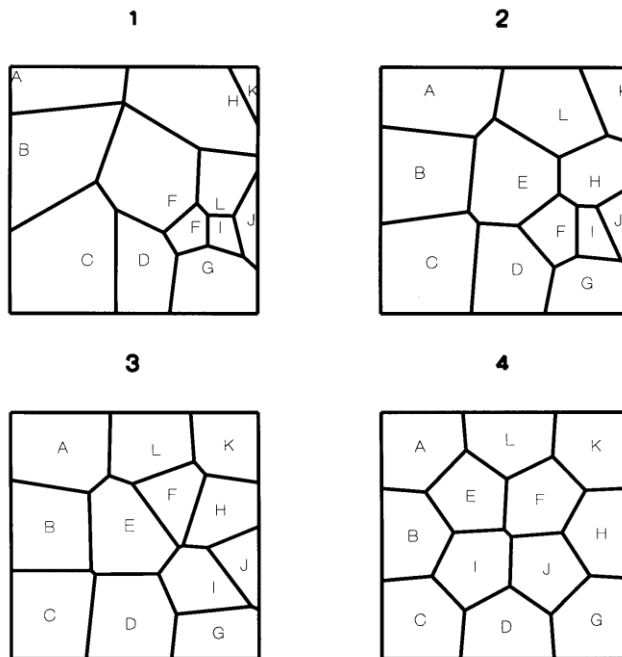
Figure 6.7 An illustration of the SADIE approach to measuring the distance to regularity for a spatial map of 12 sycamore aphids. (1) The starting arrangement of the 12 aphids with the Dirichlet tessellation outlined. (2) The first movement of the SADIE algorithm in which each aphid has been moved by the computer in the direction of a uniform pattern. (3) The position of the aphids after 4 moves. (4) The final position of the 12 aphids after 325 moves in the computer, showing a regular pattern. The sum of the moves between the initial and final positions of each aphid gives the distance to regularity measure. (Modified from Perry 1995b).

It is important to remember when sampling with quadrats that *the index of dispersion depends on quadrat size and shape*. Elliott (1977) shows this grphically for a population with a clumped pattern in which the clumps themselves are uniformly spaced (Figure 6.8). The only resolution to this problem is to sample the population with a series of quadrats of varying size and to plot empirically how the index of dispersion changes with quadrat area.

If a population is completely random in its spatial pattern within the study area, there will be no change in the index of dispersion with quadrat size, contrary to the results shown in Figure 6.8. In this special case, quadrat counts are always described by the Poisson distribution, no matter what the quadrat size or shape. It seems clear that such ideal randomness is rarely found in nature.
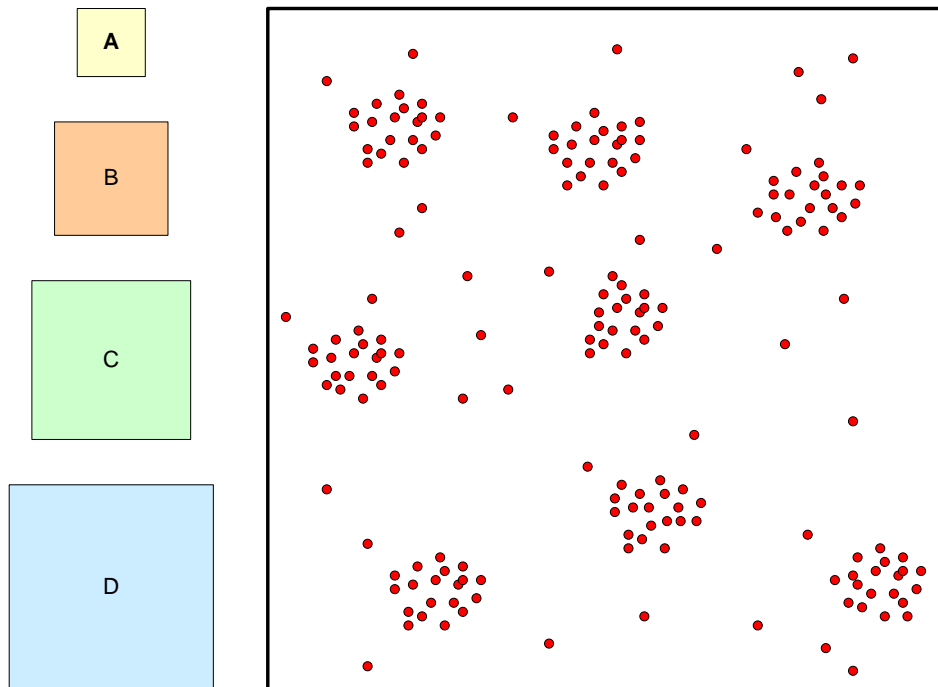
Figure 6.8 A hypothetical clumped population with regularly distributed clumps. If four quadrat sizes are used to sample this type of population, the index of dispersion obtained will show apparent randomness with quadrat size *A*, a clumped pattern with quadrat size *B*, apparent randomness with quadrat size *C*, and finally a uniform pattern with quadrat size *D*. (Modified from Elliott, 1977).

There is little information on the sampling distributions of any of the indices of dispersion, so that it is impossible to suggest any simple way of producing confidence intervals for these indices. In principle it should be possible to do this with jackknife or bootstrap techniques, but this does not seem to be discussed any of the literature on indices of dispersion.

The analysis of spatial patterns in plant and animal populations has blossomed in complexity over the past 30 years, so that many of the methods used are complex mathematically and are best implemented in computer programs. As pointed out clearly in their introduction to spatial analysis, Fortin and Dale (2005) remind ecologists that the motivation for doing spatial analysis is to detect patterns in plant and animal populations, but this is just the beginning of trying to understand the processes that cause the observed patterns.

6.5 SUMMARY

Spatial pattern is of interest in itself because we would like to understand the ecological processes determining the locations of individuals which are rarely spread at random over the landscape. Spatial pattern depends on scale, and there is no single spatial pattern that is universal even for one species. When a spatial map is available for a particular population, nearest-neighbor distances can be measured for the entire population to assess spatial pattern. Distances to second and third nearest-neighbors can provide additional information about the spatial positioning of individuals. A series of techniques of increasing sophistication are available for analyzing pattern in spatial maps.

Plant ecologists use contiguous quadrats to analyze the spatial patterning of plant populations. The variance among quadrats changes as adjoining quadrats are grouped into larger and larger units, and this change can help identify the scale of clump sizes in aggregated distributions.

For large area sampled with distance measures spatial pattern can be analyzed from data on distances from random points to nearest organisms. The degree of non-randomness can be measured with two different measures of dispersion for distance methods.

Many indices of dispersion for quadrat sampling have been proposed. For quadrat sampling the most useful one seems to be the Morisita index. New computer-intensive methods (SADIE) have also been suggested for measuring aggregation in organisms sampled with quadrats. With quadrat sampling it is critical to remember that the spatial pattern obtained usually depends on quadrat size and shape, and a series of quadrat sizes should be used to measure the scale of the patterns shown by the population. The key at the end is to understand why particular spatial patterns occur.

SELECTED READINGS

Clark, P.J. & Evans, F.C. 1954. Distance to nearest neighbor as a measure of spatial relationships in populations. *Ecology* 35: 445-453.

Dale, M. 1997. *Spatial* Pattern *Analysis in Plant Ecology*. Cambridge University Press, Cambridge.

Diggle, P.J. 1983. *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.

Engeman, R.M., Sugihara, R.T., Pank, L.F. & Dusenberry, W.E. 1994. A comparison of plotless density estimators using Monte Carlo simulation. *Ecology* **75:** 1769-1779.

Fortin, M.-J., and Dale, M.R.T. 2005. *Spatial Analysis: A Guide for Ecologists*. Cambridge University Press, Cambridge.

Hurlbert, S.H. 1990. Spatial distribution of the montane unicorn. *Oikos* **58**: 257-271.

Mumme, R.L., Koenig, W.D., & Pitelka, F.A. 1983. Are acorn woodpecker territories aggregated? *Ecology* 64: 1305-1307.

Perry, J.N. 1995. Spatial analysis by distance indices. *Journal of Animal Ecology* **64**: 303-314.

Scheldeman, X., and van Zonneveld, M. 2010. *Training Manual on Spatial Analysis of Plant Diversity and Distribution.* Biodiversity International, Rome. 179 pp.

QUESTIONS AND PROBLEMS

6.1. A series of 64 contiguous quadrats of 1 cm^3 of forest soil was counted for collembolans with the following results:

4	3	10	5	8	3	2	8	0	0	5	4	13	17	7	4
1	7	11	6	5	9	8	2	3	2	6	19	6	4	5	2
2	5	5	7	16	8	9	9	3	3	2	9	11	5	2	0
0	2	0	3	5	4	2	1	0	4	11	2	3	10	3	2

What spatial pattern do these collembolans show in this forest soil?

6.2. For a goldenrod (*Solidago* spp.) plot of 136 by 176 ft. (23,936 sq. ft.), Clark and Evans (1954) counted 89 plants and measured nearest-neighbor distances for each. They obtained

$$n = 89, \quad \sum r = 530.24, \quad \sum r^2 = 4751.5652.$$

(a) Calculate the Clark and Evans test for a random pattern for these data, and compare the results with that of Thompson's Test.

(b) Assume that no boundary strip was included in these data and calculate the Donnelly modification of the Clark and Evans test. How do your conclusions differ?

6.3. Burgess *et al.* (1982) tested the spatial patterning of acorn woodpecker territories in California by measuring the distances to the first-to-fifth nearest neighbors in

their study area. They used the woodpeckers' storage trees (granaries) as the point from which distances were measured, and got these results: (study area = 1313.5 sq. distance units):

Bird #	Nearest-neighbor distance[a]				
	1st	2nd	3rd	4th	5th
1	2.9	4.7	7.8	6.9	7.0
2	3.2	3.6	3.7	4.7	5.2
3	2.7	3.6	4.4	4.9	6.7
4	3.6	3.8	4.4	4.6	4.9
5	2.7	3.6	5.1	5.8	6.9
6	4.9	6.7	8.2	8.4	8.9
7	0.6	3.2	4.1	4.8	5.5
8	3.7	5.6	7.1	8.6	9.8
9	4.7	4.8	6.9	7.1	8.2
10	4.9	5.0	6.1	6.9	7.2
11	0.6	3.7	3.9	4.8	4.9
12	5.3	5.7	6.7	6.9	9.1
13	1.1	3.7	4.8	4.8	6.0
14	4.4	5.8	7.7	9.0	10.8
15	2.9	4.1	4.8	5.0	5.1
16	4.8	5.1	5.3	5.7	8.0
17	1.1	3.9	4.1	4.5	5.1
18	3.9	5.6	8.2	[b]	[b]
19	5.7	8.9	9.8	[b]	[b]
20	3.7	3.9	[b]	[b]	[b]
Mean distance =	3.37	4.75	5.79	6.08	7.02
Sum of squares =	273.42	487.26	694.39	666.52	892.61
Density =	0.0152	0..0152	0.0145	0.0129	0.0129

[a] Each sample distance unit = 48 m.

[b] These neighbors fell outside the sample perimeter.

Use Thompson's test to evaluate the null hypothesis of a random pattern for each of the 1st to 5th nearest neighbors. Review the controversy arising from this study (Mumme *et al.* 1983, Burgess 1983) and suggest techniques for further analysis of the problems discussed.

6.4. Neil Gilbert sampled two species on randomly selected cabbage leaves and obtained the following data:

Leaf number	Cabbage aphid	Predatory beetle
1	5	0
2	4	0
3	5	0
4	1	0
5	2	1
6	1	0
7	0	2
8	1	1
9	2	0
10	4	1
11	4	0
12	0	0
13	1	1
14	1	1
15	0	1
16	2	0
17	1	0
18	3	0
19	4	0
20	0	1
21	1	0
22	1	1
23	2	1
24	1	2

Calculate an index of dispersion for each of these two species and estimate the confidence limits for these indices.

6.5 Discuss the statistical and ecological reasons why the hypothesis of a uniform spatial pattern might not be accepted for data like that given for ants in Box 6.1 (page 239).

6.6. Calculate an index of dispersion for the data on subalpine fir trees in Box 5.2 (page 221).