7

Exploring spatial point patterns

7.1 Introduction

The simplest form of spatial data is a spatial point pattern, although approaches to their analysis are no simpler than is the case with data that have attributes (i.e. have numeric values attached to them). A spatial point pattern is simply a set of locations that correspond to events. For example, a set of points that shows the locations of trees or of people with a particular disease is a point pattern. Many tools exist for analysing point patterns. Such approaches allow assessment of, for example, the degree to which point events are clustered or dispersed. Synthetic examples of clustered and dispersed point patterns are given in Figure 7.1 (hereafter referred to as 'Point Pattern 1' or 'PP1') and Figure 7.2 ('PP2'). The concern is often to consider if the point pattern is spatially structured in some particular way—if the points are clustered then this suggests that events are more likely to occur in some places than in others. Conversely,

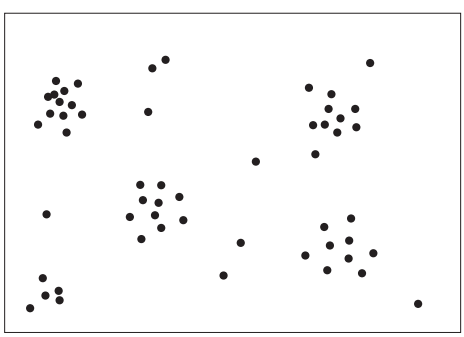

Figure 7.1 Clustered point pattern (PP1).

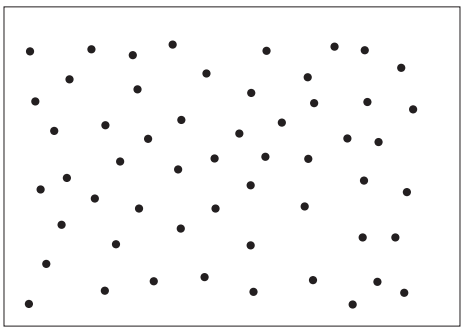

Figure 7.2 Dispersed point pattern (PP2).

if the points are dispersed, this suggests that events are likely to occur as far away as possible from other events. This chapter introduces a variety of ways for assessing and characterizing spatial point patterns. A set of points that also have values attached to them is called a marked point pattern, but the concern in this chapter is with points that have no attached attribute data. While much of the early development of approaches was in an ecological context, applications areas are now extensive. Typical applications of point pattern analysis include the exploration of clustering in disease events (an interesting study by Openshaw *et al.* (1993) notes that cases of some forms of cancer tend to cluster while others do not) and clustering in particular species of tree (relevant references appear at the end of the chapter).

Point pattern analysis can be divided into two sets of approaches: those that deal with first-order effects and those that deal with second-order effects. First-order effects are referred to in terms of intensity—that is, the mean number of events per unit area at a given location. Second-order effects, or spatial dependence, refer to the relationship between paired events in the study region (Bailey and Gatrell, 1995). So, the first refers to the number of events in an area while the second refers to structure in the point pattern. As an example of the latter, if the number of events in areas separated by a fixed distance is consistently similar for all locations and similarity in the number of events in two areas decreases as distance between these locations increases then there is spatial dependence in the point pattern at a variety of scales. First-order effects are considered in Section 7.3, while second-order effects are the subject of Section 7.4. First- and second-order effects are a function of spatial scale. In the former case the mean intensity may change smoothly from place to place over a large area. In the latter case, features with a finer scale are the concern (Bailey and Gatrell, 1995). In practical terms, it is often difficult to separate first- and second-order effects (O'Sullivan and Unwin, 2002). This chapter does not delve further into the particular problem of distinguishing between the two, but simply presents a set of tools for the analysis of point patterns and provides some pointers for their use.

Before proceeding, it is worth making one comment of note. If, for example, a point pattern represents disease events, there may be clusters of events in some places simply because there are more people in an area, for example an urban area. If we want to explore spatial clustering, it is necessary to account for the total population, often

referred to as the 'population at risk'. Methods exist that control for the population at risk but this chapter focuses on the standard unmodified measures. Nevertheless, it is an issue that should be taken into account when considering the approaches presented.

7.2 Basic measures

Analysis of spatial point patterns is, like any analysis of any form of spatial data, likely to begin with visual inspection. Again, as with any form of spatial analysis, a second step may be to compute one of a variety of descriptive measures. Two quite widely used summary measures are the mean centre and the standard distance. The mean centre of a point pattern indicates the central tendency of the points. It is simply the mean of the x and y coordinates:

$$\bar{\mathbf{x}} = (\mu_x, \mu_y) = \left(\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n} y_i}{n} \right) \tag{7.1}$$

The bold letter $\mathbf{x}$ indicates a vector representing location (i.e. x and y coordinates) and the bar above it indicates that it is the mean average value: μ_x is the mean of the x values and μ_x is the mean of the y values. The number of events is given by n. Dispersion around the mean centre can be measured with the standard distance, d_s:

$$d_s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu_x)^2 + (y_i - \mu_y)^2}{n}} \tag{7.2}$$

The mean centres and standard distances for the two point patterns illustrated in Figures 7.1 (PP1) and 7.2 (PP2) are given in Figures 7.3 and 7.4.

The mean centres and standard distances for the two examples are similar with respect to the common boundary box. However, the characteristics of the two point patterns are very different in other ways. For example, the intensity of points in Figure 7.1 is

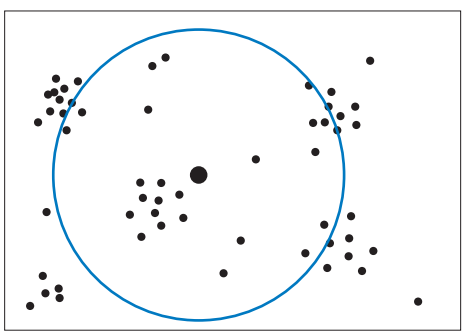

Figure 7.3 PP1: mean centre (large point) and standard distance (circle).

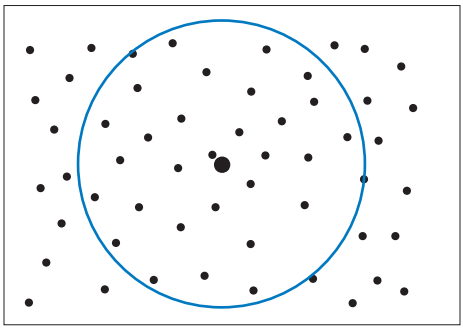

Figure 7.4 PP2: mean centre (large point) and standard distance (circle).

variable across the study region (as noted above, it is a clustered point pattern) while the intensity of points in Figure 7.2 does not vary markedly from place to place, it is a fairly regular point pattern. The rest of this chapter will focus on methods for exploring the degree to which a point pattern is clustered either globally (i.e. across the whole study area) or locally (i.e. in regions within the whole study area).

7.3 Exploring spatial variations in point intensity

A key concern in point pattern analysis is exploration of spatial variation in point pattern intensity. The following sections explore some ways of mapping event intensity. Such methods are used for exploring first-order effects (as defined in Section 7.1). The first focus is on quadrat analysis and the second focus is on a more sophisticated means of assessing spatial variation in event intensity, kernel estimation.

7.3.1 Quadrats

A common way of exploring spatial patterning in the number of events is quadrat analysis. At its simplest level, this entails superimposing a regular grid over the point pattern and counting how many points fall within each grid square. This is illustrated in Figures 7.5 and 7.6 for the point patterns shown earlier. Clearly, varying quadrat size will have an impact on analyses, as will aggregation at different levels in other applications (recall that Section 4.9 dealt with such effects). In addition, the origin of the quadrat grid (its smallest x and y coordinates) will also influence the results.

In Figure 7.5 (PP1), there are several counts that are greater than 2. In contrast, in Figure 7.6 (PP2), there are no counts greater than 2. These values reflect the clustering and regularity, respectively, in the two point patterns. There is a range of ways of assessing the degree of clustering or regularity. One possible approach is to use a measure of spatial autocorrelation (such as Moran's I, as defined in Section 4.8) to assess the degree of spatial dependence in quadrat counts. For rook's case contiguity (see Section 4.8), I for the quadrat counts in Figure 7.5 (with zeros placed in the empty cells) was 0.0017

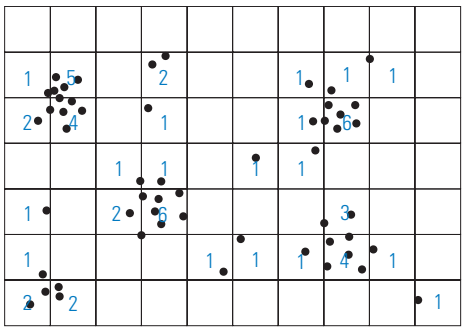

Figure 7.5 PP1: quadrats.

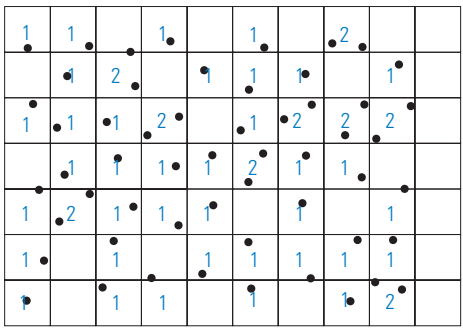

Figure 7.6 PP2: quadrats.

while for queen's case (again, see Section 4.8) it was 0.0887. For rook's case contiguity, I for the quadrat counts in Figure 7.6 was –0.0395 while for queen's case it was –0.1053. For the quadrats used in the example, PP1 appears clustered, i.e. it is positively spatially autocorrelated (albeit not to a marked degree), while PP2 appears regular or dispersed, i.e. it is negatively autocorrelated, thus confirming the impression gained from visual inspection. The I coefficient was computed using the GeoDa software (Anselin *et al.*, 2006). The value of I is, of course, partly a function of the size of the quadrats.

A simple means of assessing the degree of clustering or regularity in a point pattern is the variance/mean ratio (VMR). The VMR indicates the degree to which a point pattern departs from that predicted by the Poisson distribution, which is often used to express the probability of events occurring in a given area. The Poisson distribution can be given by:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{7.3}$$

where λ is the mean average intensity of the point pattern, k is the number of events, and e is the base of the natural logarithm (=2.71828…). In words, the equation

gives the predicted fraction of quadrats containing k events for a mean intensity λ (O'Sullivan and Unwin (2002) provide another account). In Figures 7.5 and 7.6 there are 70 quadrats and 55 events. In both cases, the mean quadrat count (the mean intensity) is $^{55}/_{70}=0.786$. As an example, the probability of there being three events in a quadrat is given by:

$$P(3)=\frac{0.786^3 e^{-0.786}}{3!}=\frac{0.485588\times 0.455664}{3\times 2\times 1}=\frac{0.221265}{6}=0.036877$$

Events like the number of points in quadrats can be modelled using the Poisson distribution and a point pattern with a Poisson distribution may be described as confirming to complete spatial randomness (CSR)—that is, there is no apparent structure. Table 7.1 gives quadrat counts (given by #) for PP1 and PP2, the counts expressed as proportions and $P(k)$. Visual inspection of the two sets of fractions suggests that the fractions for PP2 are much closer to those predicted given the Poisson distribution than are those for PP1.

Following on from the discussion above, the VMR is expected to have a value of 1 if the distribution is Poisson—a value greater than 1 indicates clustering while a value of less than 1 indicates a dispersed point pattern. Recall that, for the examples in Figures 7.5 and 7.6, there are 70 quadrats and 55 events, giving a mean quadrat count of $^{55}/_{70}=0.786$ in both cases. For the point pattern in Figure 7.5 (PP1), the variance (i.e. the mean squared difference between each quadrat count and the mean quadrat count) is $^{131.786}/_{70}=1.883$. For the point pattern in Figure 7.6 (PP2), the variance is $^{29.786}/_{70}=0.426$. The variance/mean ratios are then given by:

$$\frac{1.883}{0.786}=2.396$$

$$\frac{0.426}{0.786}=0.542$$

Table 7.1 Quadrat counts for PP1 and PP2, fractions and $P(k)$

k	PP1#	PP2#	PP1 fraction	PP2 fraction	$P(k)$
0	42	24	0.600	0.343	0.45566381
1	17	37	0.243	0.529	0.35815175
2	5	9	0.071	0.129	0.14075364
3	1	0	0.014	0.000	0.03687745
4	2	0	0.029	0.000	0.00724642
5	1	0	0.014	0.000	0.00113914
6	2	0	0.029	0.000	0.00014923
7	0	0	0.000	0.000	0.00001676
8	0	0	0.000	0.000	0.00000165
9	0	0	0.000	0.000	0.00000014
10	0	0	0.000	0.000	0.00000001

This supports the notion of PP1 exhibiting clustering (as the value is greater than 1), while visual inspection of PP2 suggests dispersal (with a value of less than 1). This supports the impression gained from inspection of Table 7.1. In addition, the analysis using Moran's *I* coefficient, summarized above, reached similar conclusions. Yet another way of assessing the degree to which a point pattern conforms to CSR is to conduct a chi-square (χ^2) test (O'Sullivan and Unwin, 2002). This is given by the variance (without division by n) of the quadrat counts divided by the mean:

$$\chi^2 = \frac{\sum (k - \mu_k)^2}{\mu_k} \tag{7.4}$$

where μ_k is the mean quadrat count (the estimate of λ). The calculations for the variances of PP1 and PP2 are given in Table 7.2.

Note that k and μ are the same for PP1 and PP2. The χ^2 statistic for PP1 and PP2 is then calculated:

$$\text{PP1: } \chi^2 = \frac{131.786}{0.786} = 167.666$$

$$\text{PP2: } \chi^2 = \frac{29.786}{0.786} = 37.895$$

The number of degrees of freedom (see Section 3.2) is given by the number of quadrats (70) minus 1, i.e. 69. Examination of a table of critical values of χ^2 (e.g. see Ebdon (1985), Appendix C) indicates that the value for PP1 is significant to greater than the 0.001 level, while the value for PP2 is not significant even at the 0.1 level. In other words, in the case of PP1 we can reject the null hypothesis that the point pattern was generated by CSR. In the case of PP2 we cannot reject the null hypothesis (see Section 3.4 for a discussion on a related topic).

The following section builds on the account of geographical weighting schemes in Section 4.7, but in a point pattern analysis context.

Table 7.2 Quadrat counts for PP1 and PP2: derivation of the variances

k	PP1#	PP2#	PP1 and PP2 $(k-\mu)$	PP1 and PP2 $(k-\mu)^2$	PP1 $\#(k-\mu)^2$	PP2 $\#(k-\mu)^2$
0	42	24	–0.786	0.617796	25.947432	14.827104
1	17	37	0.214	0.045796	0.778532	1.694452
2	5	9	1.214	1.473796	7.368980	13.264164
3	1	0	2.214	4.901796	4.901796	0.000000
4	2	0	3.214	10.329796	20.659592	0.000000
5	1	0	4.214	17.757796	17.757796	0.000000
6	2	0	5.214	27.185796	54.371592	0.000000
Total	70	70			131.785720	29.785720

7.3.2 Kernel estimation

Kernel estimation (KE) is a more sophisticated means of exploring spatial variations in event intensity. Such approaches are often used in the identification of clustering 'hot spots'. If we want to estimate the intensity of points over an area we can simply calculate the number of events within a radius around the nodes of a grid and divide this amount by the area concerned. In other words, we in effect overlay onto the point pattern a set of points on a regular grid and compute the event intensity within the neighbourhood of each point on the grid. This is called the naïve estimator. The naïve intensity estimate is given by:

$$\hat{\lambda}(\mathbf{x}) = \frac{\#(C(\mathbf{x},d))}{\pi d^2} \tag{7.5}$$

where $\#(C(\mathbf{x},d))$ indicates the number (#) of events in the circle $C(\mathbf{x},d)$ that has as its centre the location $\mathbf{x}$, and has the radius d, with its area given by πd^2. The bold lower case $\mathbf{x}$ is matrix vector notation as introduced in Section 3.3. In this case, $\mathbf{x}$ refers to a location with coordinates x,y.

Figure 7.7 shows intensity estimates for PP1 using radii of 25 and 50 units. The differences in minimum and maximum intensities and in the spatial patterning are immediately apparent. The maximum intensity is greater for a radius of 25 units as this smaller radius 'picks out' the small clusters. Their effect is diminished when the larger radius (50 units) is used. The two groups of clusters notable when the 25 unit radius is used become, in effect, merged into one as the size of the radius is increased to 50 units.

KE can be expanded to make use of a geographical weighting scheme (a kernel function) whereby the influence of the points varies depending on how far they are from the centre of the window. The general idea of geographical weighting was

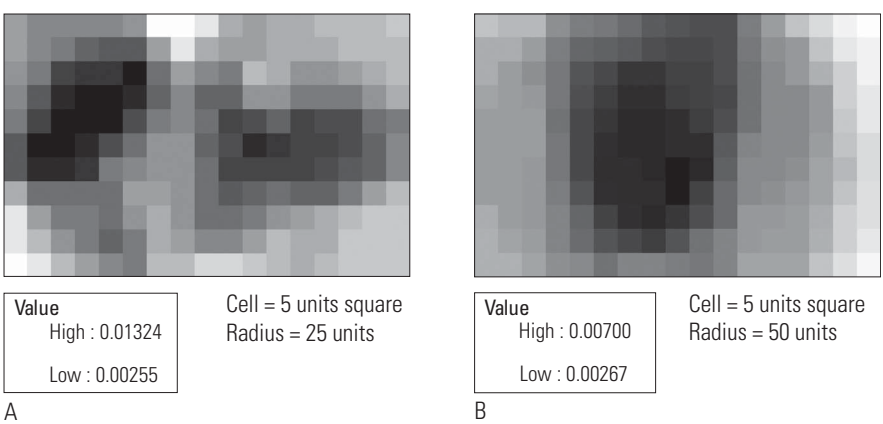

Figure 7.7 Intensity estimates for PP1 using radii of 25 units (A) and 50 units (B). The values are events per square unit.

explored in Section 4.7, but is recapped here to illustrate KE. The KE of intensity is given by:

$$\hat{\lambda}_k(\mathbf{x}) = \sum_{i=1}^{n} \frac{1}{\tau^2} k\left(\frac{\mathbf{x} - \mathbf{x}_i}{\tau}\right) \tag{7.6}$$

where τ is the bandwidth (determining the size of the kernel, see Section 8.4 for a related discussion) and $\mathbf{x} - \mathbf{x}_i$ indicates the distance between the centre of the kernel ($\mathbf{x}$) and the location $\mathbf{x}_i$. A kernel function is used to give larger weights to nearby events (points) than events that are more distant.

There is a variety of different kernels that have been used for KE. The quartic kernel is encountered frequently in the point pattern analysis literature (Bailey and Gatrell, 1995):

$$k(\mathbf{u}) = \begin{cases} \frac{3}{\pi}(1 - \mathbf{u}^T\mathbf{u})^2 & \text{for } \mathbf{u}^T\mathbf{u} \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{7.7}$$

where $\mathbf{u}$ is d_i/τ and d_i is the distance from the centre of the kernel, and superscript T indicates the transpose of the matrix (see Appendix A for a definition). In short, values are summed for all cases where the distance is less than or equal to the bandwidth. The quartic kernel is illustrated in Figure 7.8 and a section through the quartic kernel is shown in Figure 7.9. The KE with the quartic kernel can be given by (Bailey and Gatrell, 1995):

$$\hat{\lambda}_k(\mathbf{x}) = \sum_{d_i \le \tau} \frac{3}{\pi\tau^2}\left(1 - \frac{d_i^2}{\tau^2}\right)^2 \tag{7.8}$$

The application of this equation is outlined below.

The naïve estimator and a kernel estimator are illustrated with reference to two sets of points, labelled A and B, in Figure 7.10. We will take each set as a subset (contained

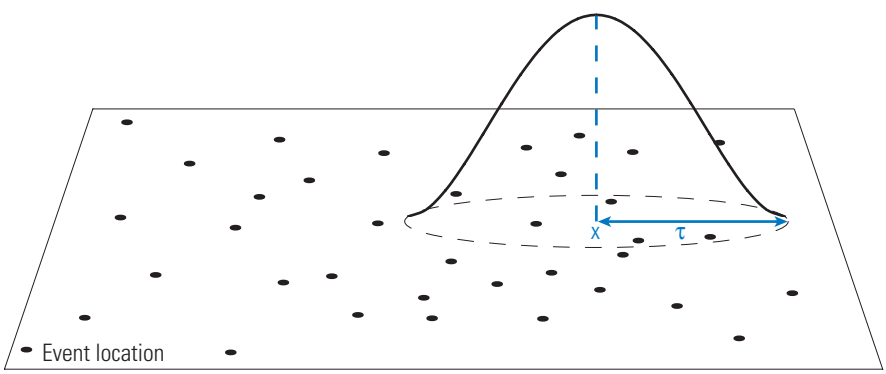

Figure 7.8 Quartic weighting scheme, with bandwidth τ.

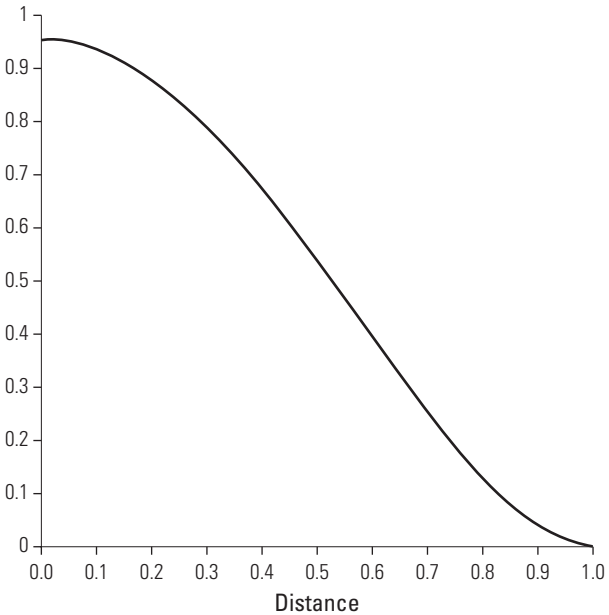

Figure 7.9 Quartic kernel section.

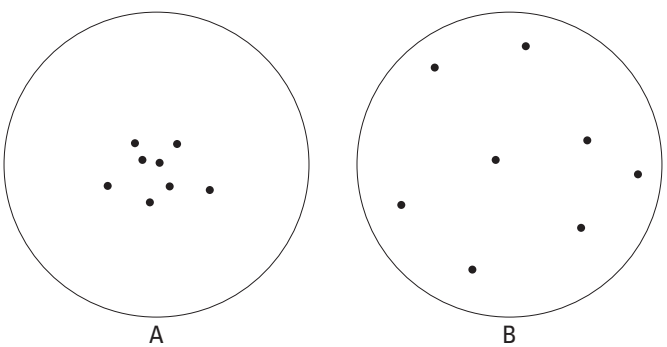

Figure 7.10 Points within a moving window with a radius of 25 units.

within the circles given, which represent a moving window with a radius of 25 units) of a larger point pattern. In both cases, there are eight points.

Using the naïve estimator, both point patterns A and B have the same intensity—the area of the circle is given by $\pi\tau^2 = 3.14159 \times 25^2 = 1963.495$. Given Equation 7.5, in both cases the intensity estimate is given by $^{8}/_{1963.495} = 0.004074$. Using a kernel estimator, point pattern A has a greater intensity than point pattern B.

Kernel estimation is now illustrated using the same example. Working through Equation 7.8, the left-hand side (LHS) of the equation, $^{3}/_{\pi\tau^2}$, is computed first. The bandwidth, τ, is set at 25 units, so $^{3}/_{\pi\tau^2} = {}^{3}/_{(3.14159 \times 25^2)} = 0.00152$.

The right-hand side (RHS) of Equation 7.8 is $(1-d_i^2/\tau^2)^2$. Taking one distance value as an example (0.927) this leads to:

$$\left(1-\frac{0.927^2}{25^2}\right)^2=\left(1-\frac{0.85933}{625}\right)^2=0.99863^2=0.99726$$

Tables 7.3 (for point pattern A) and 7.4 (for point pattern B) show the *x* and *y* coordinates of each event, the distance of the event from the window centre, and computed values of the RHS of Equation 7.8 followed by the LHS of the equation (which has a value of 0.00152 for all cases given a bandwidth of 25 units) multiplied by the RHS. The sum of products in this last column is also given.

The difference in the estimated intensities for the two methods will decrease as the size of the bandwidth increases. This is indicated by Figure 7.11, which shows the KE values for bandwidths of 5, 10, 15, 20, 25, and 30 units. Note how for point pattern A, which is visibly more clustered around the centre of the window, the KE value is much larger for a bandwidth of five units than is the value for point pattern B, but this difference decreases as the size of the kernel bandwidth is increased.

Table 7.3 Point pattern A: intensity estimate using the quartic kernel with a bandwidth of 25 units (distance is from the centre of the kernel)

x	*y*	Distance	RHS	LHS×RHS
–16.670	–2.156	0.927	0.997	0.00152
–19.794	–1.079	2.378	0.982	0.00150
–15.377	–5.496	4.235	0.943	0.00144
–14.084	1.507	4.832	0.927	0.00142
–21.087	1.615	4.960	0.923	0.00141
–18.717	–8.082	6.330	0.876	0.00134
–25.613	–5.389	8.802	0.767	0.00117
–8.912	–6.035	9.592	0.727	0.00111
			Sum	0.01091

Table 7.4 Point pattern B: intensity estimate using the quartic kernel with a bandwidth of 25 units (distance is from the centre of the kernel)

x	*y*	Distance	RHS	LHS×RHS
–19.687	–1.079	2.276	0.983	0.00150
–4.603	2.153	13.554	0.499	0.00076
–5.572	–12.392	15.948	0.352	0.00054
–23.565	–19.180	18.334	0.214	0.00033
–35.309	–8.513	18.966	0.180	0.00028
–14.731	17.776	19.838	0.137	0.00021
–29.815	14.220	20.225	0.119	0.00018
3.801	–3.449	21.409	0.071	0.00011
			Sum	0.00390

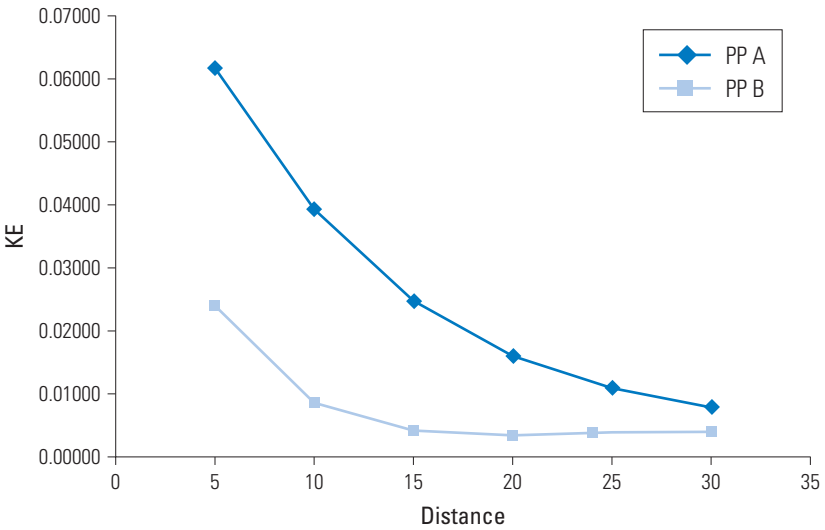

Figure 7.11 KE for point patterns A and B for different bandwidths.

Clearly, the size of the kernel bandwidth will have a marked impact on the intensity estimates. In practice, a variety of bandwidths is often used and the difference in results assessed. In the cases described above, the size of the kernel bandwidth is the same at all locations, but techniques exist such that the bandwidth can be varied from place to place as a function of local event intensity, with a small bandwidth in areas with intense point patterns and a large bandwidth in areas with a low event intensity (Bailey and Gatrell, 1995; Brunsdon, 1995).

7.4 Measures based on distances between events

The second-order properties of a point pattern (as defined in Section 7.1) can be measured using information on distances between events. This section deals with such approaches. The initial concern is with methods that make use of information on the nearest neighbours of events. The second concern is with the *K* function, a widely used means of assessing the degree of spatial dependence in point patterns (i.e. does the number of points in one area tend to be similar to the number of points in another area some fixed distance away?)

7.4.1 Nearest-neighbour methods

Various measures exist that are based on the nearest neighbour to each event. The mean nearest-neighbour distance is one such measure and it is given by:

$$\bar{d}_{\min} = \frac{\sum_{i=1}^{n} d_{\min}(\mathbf{x}_i)}{n} \tag{7.9}$$

where n is the number of events and $d_{\min}$ is the distance from the nearest event. In words, the distance of each event i at location $\mathbf{x}_i$ to its nearest neighbour and the average of these distances are computed. These statistics are used to compute two other measures called the G and F functions, which will be described next.

The G and F functions allow the exploration of event to event nearest-neighbour distances. The G function is defined as the cumulative frequency distribution of the nearest-neighbour distances. It is given by:

$$G(d) = \frac{\#(d_{\min}(\mathbf{x}_i) < d)}{n} \tag{7.10}$$

In words, $G(d)$ gives the proportion (since the count is divided by n) of nearest-neighbour distances that are less than distance d. Obtaining G for different values of d enables assessment of the degree of clustering at different spatial scales. For example, for a clustered point pattern, G will increase markedly as distance increases for small distances. For a regular point pattern the increase will be more gradual. Figure 7.12 shows the G function for the point patterns in Figures 7.1 (PP1) and 7.2 (PP2) for distance steps of 2.5 units. Clearly, the values of G for PP1 are larger at smaller distances than they are for PP2, indicating the greater degree of clustering in PP1 than in PP2. A value of 1 for G corresponds to distances that are larger than the maximum nearest-neighbour distance for a given point pattern. In other words, for PP1, no nearest-neighbour distance is greater than 15 units, while for PP2 no nearest-neighbour distance is greater than 12.5 units.

The F function is similar to the G function, but instead of the events a sample of point locations is selected randomly from anywhere in the study area—that is,

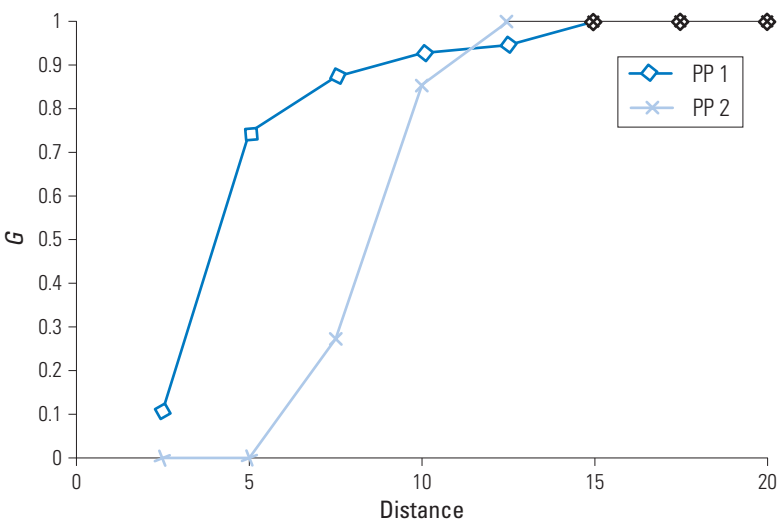

Figure 7.12 *G* function for the point patterns in Figures 7.1 (PP1) and 7.2 (PP2).

the nearest-neighbour distances are computed for randomly selected locations and not for the point event locations. The *F* function is given by:

$$F(d) = \frac{\#(d_{\min}(\mathbf{x}_i, X) < d)}{m} \tag{7.11}$$

where $d_{\min}(\mathbf{x}_i, X)$ is the minimum distance from the location $\mathbf{x}_i$ in the randomly selected set of locations to the nearest event in the point pattern, X, and m is the number of randomly selected locations. Both *G* and *F* can be plotted against *d* to allow exploration of changes in clustering with distance. The *G* and *F* functions may vary, for example, for a clustered pattern, values of *G* may be larger than values of *F* for small distances as in the latter case only a random sample is used and the impact of clusters may be diminished (Bailey and Gatrell, 1995; O'Sullivan and Unwin, 2002). One benefit of the *F* function is that the sample size *m* can be varied, with larger values of *m* giving a smoother curve. Comparison of the *G* and *F* functions for the same point pattern may be informative since the two functions bring out different characteristics of the point pattern (O'Sullivan and Unwin, 2002).

The most widely used measure of spatial dependence in point patterns is the *K* function, which is the subject of the following section.

7.4.2 *K* function

Whereas the *G* and *F* functions are based on nearest neighbours, the *K* function is based on distances between all events in the region of study. The *K* function for distance *d* can be given by:

$$\hat{K}(d) = \frac{|A|}{n^2} \sum_{i=1}^{n} \#(C(\mathbf{x}_i, d)) \tag{7.12}$$

As defined previously, $\#(C(\mathbf{x}_i, d))$ indicates the number (#) of events in the circle $C(\mathbf{x}_i, d)$, which has as its centre the location $\mathbf{x}_i$ and radius *d*, and $|A|$ indicates the area of the study region. Recall that the hat (^) indicates that it is an estimate.

The *K* function can be computed by following several steps:

1. Go to an event and count the number of other events within a set radius of the event.
2. Do the same for all other events, adding the number of points within that set radius to the number of points in that radius for all events visited.
3. Once all events have been visited, the total counts within the distance band, *d*, is scaled (multiplied) by $\frac{|A|}{n^2}$.
4. Increase the radius a fixed amount and go back to step 1, repeating the process to the maximum desired radius.

As an example, for PP1 (Figure 7.1) there were five events within 5 units of the first event visited, so the unscaled version of $K(5) = 5$. There were four events within 5 units

of the second event visited, so the unscaled version of $K(5)$ becomes 9 and so on until all events have been visited.

For PP2 (Figure 7.2), the total number of events within 5 units of all events was zero (i.e. there were no other events within 5 units of any event) and for PP1 it was 122. As noted above, for each distance, the count is multiplied by $\frac{|A|}{n^2}$. The area of the study region is 6665.744 square units and there are 55 events in both PP1 and PP2. For PP1 $\hat{K}(5)$ is therefore obtained from:

$$\frac{6665.755}{55^2} \times 122 = 268.833$$

and for PP2, $\hat{K}(5)$ is:

$$\frac{6665.755}{55^2} \times 0 = 0$$

The expected value of K is given by:

$$E[K(d)] = \frac{\lambda \pi d^2}{\lambda} = \pi d^2 \tag{7.13}$$

where λ is the average intensity of the point pattern (but is not needed to calculate the expected value of K). Equation 7.13 gives the expected value for a point pattern that is said to be the outcome of a CSR process (see Section 7.3.1). A value of $K(d)$ greater than the expected value (i.e. πd^2) suggests a clustered point pattern, while a value of $K(d)$ less than the expected value suggests a regular point pattern. In practice, the L function, a transformed version of the K function, is often computed. The L function for distance d is given by:

$$\hat{L}(d) = \sqrt{\frac{\hat{K}(d)}{\pi}} - d \tag{7.14}$$

Values of the L function of less than 0 indicate regularity, while values greater than 0 suggest clustering. The K function is illustrated in Figure 7.13 using the point patterns illustrated in Figures 7.1 (PP1) and 7.2 (PP2), with the expected values ($E(K)$) also plotted. In this example, the smallest distance band was 2.5 units and this was increased in steps of 2.5 units to a maximum of 25 units. The L function is illustrated in Figure 7.14.

Note that the values of K for PP1 are clearly larger than those for PP2 and the expected value for smaller ($\leq$17.5 units) distances, while the expected values are larger than those for PP1 for distances of 20 and greater. In contrast, values for PP2 are clearly smaller than the values for both PP1 and the expected values up to a distance band of 20 units. In words, PP1, defined as a 'clustered' point pattern, is indeed clustered at short distances, while PP2, defined as a 'regular' point pattern, is dispersed at small

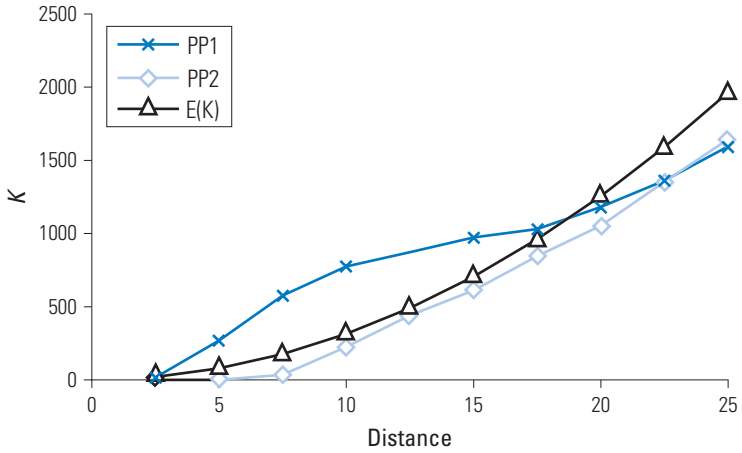

Figure 7.13 *K* function and expected values (*E*(*K*)) for the point patterns in Figures 7.1 (PP1) and 7.2 (PP2).

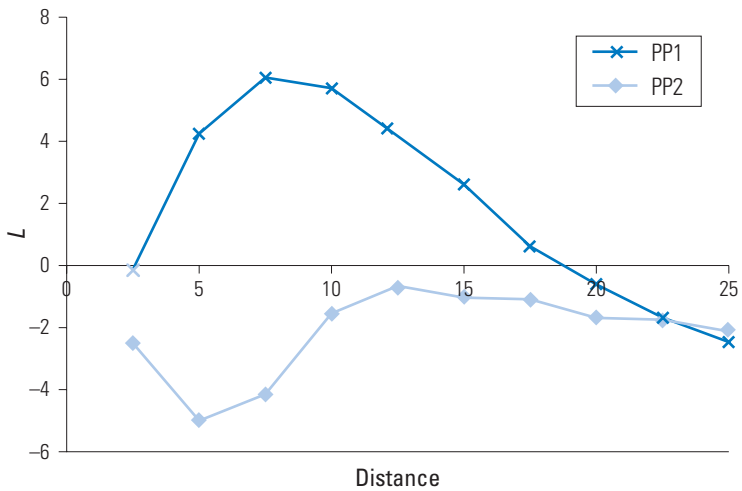

Figure 7.14 *L* function for the point patterns in Figures 7.1 (PP1) and 7.2 (PP2).

distances. In an application that makes use of the *K* function, O'Brien *et al.* (2000) demonstrate how different forms of cancers in humans and dogs tend to cluster over particular distance ranges. Such tools are a powerful means of assessing the structure of point patterns and they may provide evidence to support further analyses (e.g. if disease events cluster at particular scales we are likely to be interested in interpreting this finding).

An important issue in point pattern analyses concerns edge effects. Particularly for larger distance bands, parts of the circles drawn around points may often fall outside

of the study area. To avoid biased results, it is necessary to take this factor into account (see Bailey and Gatrell (1995) for a discussion about this issue). For example, if KE is being conducted, and part of the kernel falls outside of the study area, then it will be necessary to account for the area of the kernel that is not within the study area.

7.5 Applications and other issues

Many different properties can be represented as point patterns. These include disease events (e.g. Hill *et al.*, 2000), trees (e.g. Li and Zhang, 2007), and concentrations of minerals in rock (Lloyd, 2006). Other applications are detailed by Bailey and Gatrell (1995) and Diggle (2003). This chapter provides only a brief outline of some key techniques for the analysis of point patterns. Consideration of the population at risk (e.g. accounting for the fact that disease rates tend to be higher in urban areas as there is a greater density of people in such areas than elsewhere), edge effects, and testing of point patterns (e.g. assessing how far a point pattern is clustered) are mentioned only quite briefly. Further issues that are not covered include the extension of the analysis of spatial point patterns to include a time element, analysis of marked point patterns (i.e. points with values attached), and techniques for the identification of clusters (i.e. specific locations with clusters as opposed to clustering of point patterns in general). Another issue that is not explored is the use of Monte Carlo randomization procedures in assessing point patterns relative to complete spatial randomness. With such approaches, multiple point patterns that are CSR are generated and these can be used to assess significant departures of real point patterns from CSR. Bailey and Gatrell (1995) describe such a procedure with respect to the L function. For the L function, it is common practice to derive approximate upper and lower 5% confidence intervals from simulated values, and these values can then be plotted (see Fotheringham *et al.* (2000) for an example). The next section presents a case study that makes use of data provided on the book website.

7.6 Case study

The following case study is based on locations in England that are represented on the Gough Map of Great Britain (dating to *c.* 1360). The data were captured as part of a research project funded by the British Academy[1] (see Lloyd and Lilley, 2009). The aim of the project was to develop a digital representation of the Gough Map. The analysis entailed using kernel estimation to explore local variation in event (i.e. place) intensity. Such an analysis is useful in that the results can be compared to those obtained using other data sources (e.g. contemporary records of taxation, etc.) to help

1 http://www.qub.ac.uk/urban_mapping/gough_map/

assess how representative the selection of places on the Gough Map is. For example, is the intensity of events (i.e. places) different for different sources and does this tell us anything about the purpose of these sources or where they were compiled? The intensity map was generated using ArcGIS™ Spatial Analyst. The kernel radius was set to 25 km and the cell size was 5 km. ArcGIS uses a kernel based on that defined by Silverman (1986, Equation 4.5). This is the same form as the kernel defined in Equation 7.7. The output gave intensities in square metres and these values were multiplied by one million to give intensities in square kilometres (as $1000\,\text{m} \times 1000\,\text{m} = 1{,}000{,}000\,\text{m}^2$). The locations are shown in Figure 7.15 and the map of intensities is given in Figure 7.16.

The greater intensity in some regions (e.g. the south-east and mid- to north-east of England) is partly a function of varying numbers of places in England during the

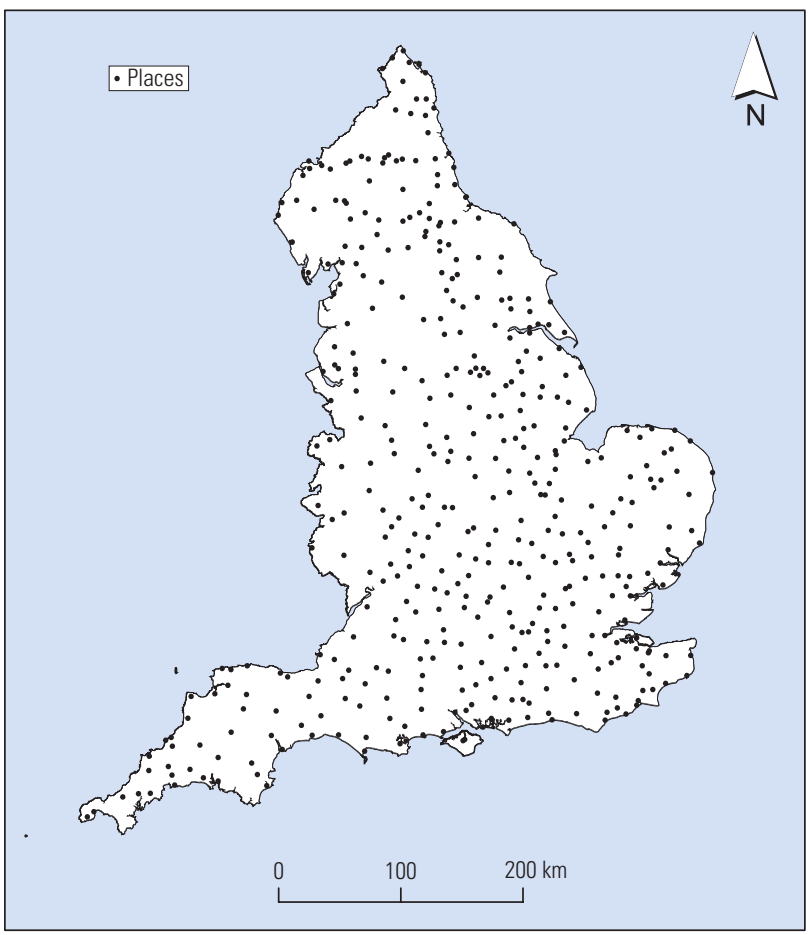

Figure 7.15 Places in England shown on the Gough Map. This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown (reproduced under the terms of the Click-Use Licence).

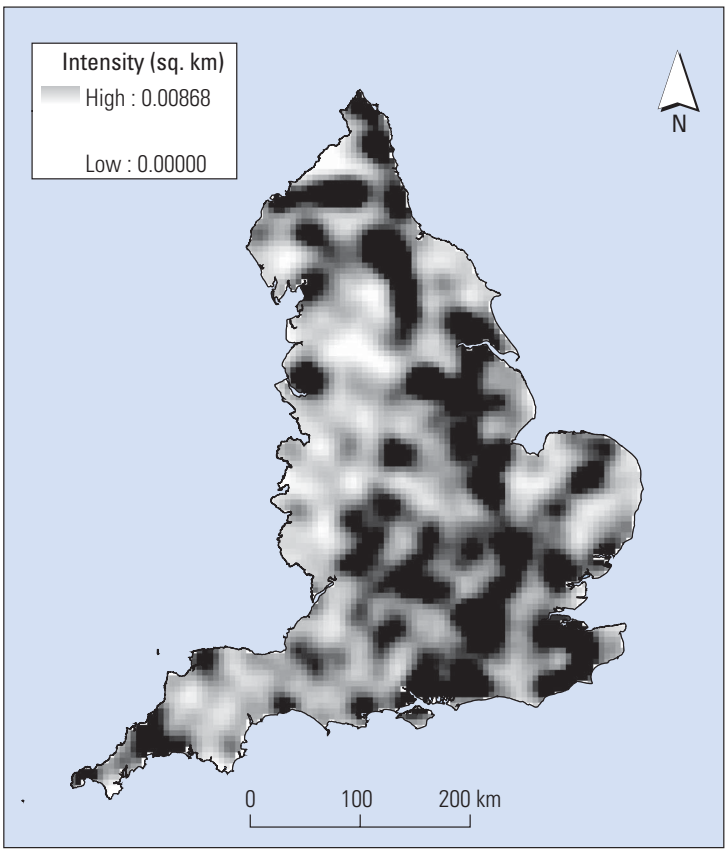

Figure 7.16 Event intensity for places in England shown on the Gough Map. This work is based on data provided through EDINA UKBORDERS with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown (reproduced under the terms of the Click-Use Licence).

period of the map's construction, but it is also partly due to biases on the part of the map's creators, who selected some places of relatively minor importance at the time, while some quite high status centres do not appear on the map.

The *K* function could also be computed to assess clustering at different spatial scales, given that there are 461 places in the data set and that the area of England is approximately 130,410 square km (i.e. 130,410,000,000 square m), there is little suggestion of clustering at any spatial scale *globally*. In terms of event intensity, there are clearly spatial variations from place to place. As such, one possible way of extending the analysis is to use a *locally* computed *K* function (see Lloyd, 2006) to explore the possibility of clustering at different scales in different regions.

The synthetic data shown in Figures 7.1 and 7.2 are provided on the book website as is a guide to how the *L* function (see Figure 7.14) can be computed using these data with freely available software.

Summary

This chapter has discussed visual examination of spatial point patterns, summary measures, measures of event intensity, measures based on distances between events (event–event distances), and point–event distances, as well as some other issues. The focus has been on the description of point patterns in terms of the degree to which they are clustered or dispersed/regular. The techniques detailed enable extraction of information on various aspects of point patterns, including spatial variation in event intensity and spatial dependence in point patterns.

Further reading

Other introductions to point pattern analysis are provided by **Bailey and Gatrell (1995)**, **Fotheringham *et al.* (2000)**, **O'Sullivan and Unwin (2002)**, **Diggle (2003)**, **Waller and Gotway (2004)**, and **Lloyd (2006)**. Each of these accounts provides information on the methods detailed in this chapter as well as information on other issues, including analysis of space–time point patterns and testing of point patterns. A good summary of some approaches to the identification of clusters is given by **Waller and Gotway (2004)**.


The next chapter also deals with analysis of spatial patterning, but with data that have attributes (e.g. values attached to them) rather than the simple point events that were the focus in this chapter.