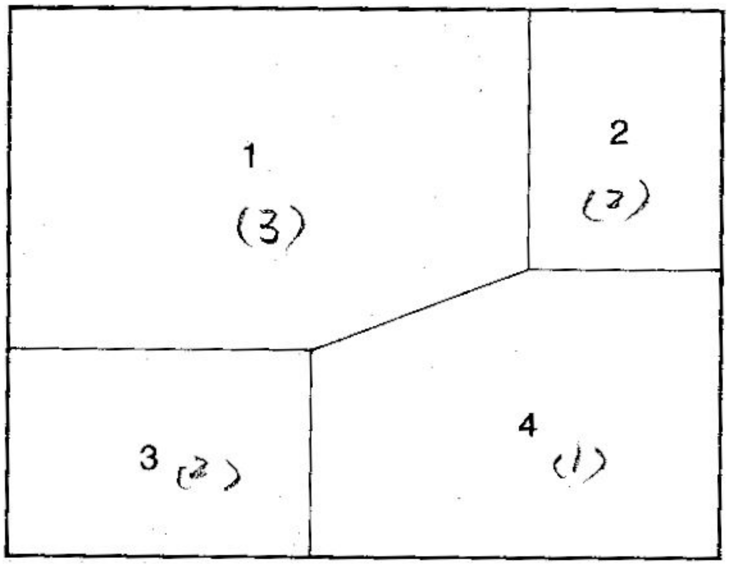

Figure 3. Example data set for calculation of the Geary and Moran coefficients.

In fact Geary arranged the index so that c would have the value of 1 when attributes are distributed independently of location, dropping below 1 when similar attributes coincide with similar locations, and above 1 otherwise. Somewhat confusingly, then, positive spatial autocorrelation corresponds to c less than 1, zero to c equal to 1, and negative to c greater than 1.

To clarify the meaning of the terms, the calculation of the Geary index can be illustrated with the simple example shown in Figure 3, using binary weights based on adjacencies.

Example calculation of the Geary index:

$n = 4 \qquad z_1 = 3, \quad z_2 = 2, \quad z_3 = 2, \quad z_4 = 1$

$$w = \begin{matrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{matrix}$$

$$\bar{z} = \sum_i z_i \,/\, n = 8/4 = 2$$

$$c_{ij} = (z_i - z_j)^2 \qquad c = \begin{matrix} 0 & 1 & 1 & 4 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 4 & 1 & 1 & 0 \end{matrix}$$

$$\sum_i \sum_j w_{ij}\, c_{ij} = 0\times0 + 1\times1 + 1\times1 + 1\times4 + 1\times1 + 0\times0 + 0\times0 + 1\times1 + 1\times1 + 0\times0 + 0\times0 + 1\times1 + 1\times4 + 1\times1 + 1\times1 + 0\times0 = 16$$

$$\sum_i \sum_j w_{ij} = 10$$

$$\sigma^2 = \sum_i (z_i - \bar{z})^2 \, / \, (n-1) = 2/3$$

$$c = \sum_i \sum_j w_{ij} \, c_{ij} \, / \, 2\sigma^2 \sum_i \sum_j w_{ij} = 16 \, / \, 2 \times (2/3) \times 10 = 1.200$$

To give a visual impression of the range of c values, Figure 4 illustrates a number of different arrangements of the same set of attributes over a map of areas. The figures are the population densities of the 51 census tracts of the city of London, Ontario, as reported in the 1971 census, and they have been distributed among the 51 tracts to show the actual pattern (4a), as well as artificial patterns of maximally positive (4b), and maximally negative (4c) autocorrelation. Note that each pattern is independent of the others. So although there is a tendency for a peak of population density to occur at the city centre in the actual data, peaks can occur anywhere in the city in the simulations.

We will leave the Geary index for now in order to continue with a brief review of the other autocorrelation measures which have been devised for the various object and attribute types. The next chapter will discuss the Geary index in more detail, and provide a number of applications.

1.3.2 Moran's index

Moran's index (Moran, 1948) provides an alternative to Geary's for the same data context, and in most applications both are equally satisfactory. Perhaps the only obvious advantage of one over the other is that the Moran index is arranged so that its extremes match our earlier intuitive notions of positive and negative correlation, whereas the Geary index uses a more confusing scale (but see the later discussion of this issue in Chapter Two). The Moran index is positive when nearby areas tend to be similar in attributes, negative when they tend to be more dissimilar than one might expect, and approximately zero when attribute values are arranged randomly and independently in space. These relationships are summarized in Table 1 below.

Table 1: Correspondence between Geary, Moran and conceptual scales of spatial autocorrelation

Conceptual	Geary c	Moran I
Similar, regionalized, smooth, clustered	$0 < c < 1$	$I > 0$*
Independent, uncorrelated, random	$c = 1$	$I < 0$*
Dissimilar, contrasting, checkerboard	$c > 1$	$I < 0$*

* The precise expectation is $-1/(n-1)$ rather than 0.

The attribute similarity measure used by the Moran index makes it analogous to a covariance between the values of a pair of objects:

$$c_{ij} = (z_i - \bar{z})(z_j - \bar{z}) \qquad (7)$$

where $\bar{z}$ denotes the mean of the attribute variable, as before.

Instead of two variables, x and y, c_{ij} measures the covariance between the value of the variable at one place and its value at another. This idea of autocovariance will appear in other contexts later in this volume.

The remaining terms in the Moran index are again designed to constrain it to a fixed range:

$$I = \sum_i \sum_j w_{ij}\, c_{ij} \,/\, s^2 \sum_i \sum_j w_{ij} \qquad (8)$$

where s^2 denotes the sample variance

$$\sum_i (z_i - \bar{z})^2 / n$$

As in the Geary index, the w_{ij} terms represent the spatial proximity of i and j and can be calculated in any suitable way. For comparison, I is also calculated below for the data shown in Figure 3. The precise relationship between the two indices will be discussed further in Chapter 2, and we will also examine the question of whether the maximum and minimum values of both indices can be known, as well as issues of sampling and hypothesis testing.

Example calculation of the Moran index:

$$c_{ij} = (z_i - \bar{z})(z_j - \bar{z}) \qquad c = \begin{matrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{matrix}$$

$$\sum_i \sum_j w_{ij}\, c_{ij} = -2$$

$$s^2 = \sum_i (z_i - \bar{z})^2 / n = 2 / 4 = 0.5$$

$$I = \sum_i \sum_j w_{ij}\, c_{ij} \,/\, s^2 \sum_i \sum_j w_{ij} = -2 / 0.5 \times 10 = -0.400$$

We have now defined two indices of spatial autocorrelation for area objects and interval attributes. The remaining sections of this chapter review a number of other measures for alternative data types.