SHORT COMMUNICATION
On the use of Lloyd’s index of patchiness

N. BEZ*
Centre de Géostatistique, 35 Rue Saint Honoré, F–77305 Fontainebleau, France

ABSTRACT
Lloyd’s index of patchiness, and to a lesser extent, of mean crowding, are often applied to (fish or ichthyo-plankton) densities, although they were designed for use with counts of individuals. The consequences of this erroneous application are described with examples. Alternative statistics are presented for application to densities. Like the mean crowding, these statistics are robust against zeroes and cannot be considered as spatial statistics.

Key words: aggregation, counts, crowding, densities, zeroes

INTRODUCTION
Ichthyoplankton analyses, among other marine biology studies, sometimes refer to the Lloyd’s index of patchiness (de Nie, 1980; Hewitt, 1981; Houde, 1985; Stabeno et al., 1996). The formula associated with this index is

\[ P = 1 + (\sigma^2 - m)/m^2 \]

where \(m\) and \(\sigma^2\) are the mean and the variance of the sample values. This formula, established by Lloyd in 1967, originally referred to a count of fish. A fish count is one case where subtracting a mean from a variance is meaningful. When fish counts have a Poisson distribution, the mean equals the variance and the index of patchiness equals 1. Indices larger than 1 are representative of overdispersed distributions, where the variance exceeds the mean.

However, in the literature, Lloyd’s index of patchiness is often computed with fish densities (i.e. numbers per unit volume or per unit area). In this case, subtracting a mean from a variance is no longer meaningful owing to the heterogeneity of the units. The results are also clearly different if they are computed on fish densities rather than fish counts.

As an example, consider a situation where we have been able to fish exhaustively some blocks (quadrats in Lloyd’s terminology) of \(S = 10\) square nautical miles (n.mi.\(^2\)). Two cases are considered: the first with two blocks (survey 1) and the second with an added third block (survey 2). Take the number of fish caught (denoted \(q_i\) with \(i \in [1,3]\)) as 80 for the first block, 20 for the second and 0 for the third, which gives the following densities (denoted \(z_i\) with \(i \in [1,3]\)): 8, 2, 0 as numbers of fish per square nautical mile (Table 1). For the first survey, the mean number of fish per block is 50 and its variance 900 so that the index of patchiness is equal to 1.34. Using fish densities instead of fish counts gives a value of 1.16.
Table 1. Presentation of the statistics: notation, formulae and values in the case of two examples. Blank cells indicate those cases where the statistics are inappropriate.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Formula</th>
<th>Survey 1</th>
<th>Survey 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total abundance</td>
<td>$Q = \Sigma q_i = S \times \Sigma z_i$</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Mean</td>
<td>$m$</td>
<td>50</td>
<td>33</td>
</tr>
<tr>
<td>Variance</td>
<td>$\sigma^2$</td>
<td>900</td>
<td>1155.6</td>
</tr>
<tr>
<td>Lloyd’s index of patchiness</td>
<td>$P = 1 + [(\sigma^2 - m) / m^2]$</td>
<td>1.34 (incorrect use)</td>
<td>2.01</td>
</tr>
<tr>
<td>Mean crowding</td>
<td>$[\Sigma(q_i - 1)q_i] / \Sigma q_i$</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Mean density per individual</td>
<td>$\Sigma z_i^2 / \Sigma z_i$</td>
<td>—</td>
<td>6.8</td>
</tr>
<tr>
<td>Index of aggregation</td>
<td>$\Sigma z_i^2 / [S \times (\Sigma z_i)^2]$</td>
<td>—</td>
<td>0.068</td>
</tr>
<tr>
<td>Mean density per individual</td>
<td>$\Sigma z_i^2 / S$</td>
<td>—</td>
<td>6.8</td>
</tr>
<tr>
<td>Index of aggregation</td>
<td>$\Sigma z_i^2 / [S \times (\Sigma z_i)^2]$</td>
<td>—</td>
<td>0.068</td>
</tr>
</tbody>
</table>

SENSITIVITY TO ZEROES: DOMAIN-FREE STATISTICS

The total number of individuals in the population, denoted $Q$, is 100 in both of the previous examples. Such a statistic, which does not depend on whether the domain chosen for the computation includes zeroes or not, is qualified here as a ‘domain-free statistic’. The mean and the variance of either the fish counts or the fish densities are modified by the inclusion of zeroes. These are not domain-free statistics.

In fact, Lloyd first defined the mean crowding as the mean number of neighbours that can be found around any single individual of a population. In our examples (Table 1), taking a neighbourhood equal to the sampling block size, we obtain the following:

- each individual present in the first block has 79 neighbours;
- each individual present in the second block has 19 neighbours;
- no neighbour can be counted in the third empty block.

So, the mean crowding given by the formula $[\Sigma(q_i - 1)q_i] / \Sigma q_i$ is in both cases: $(80 \times 79 + 20 \times 19) / (80 + 20) = 67$ number of neighbours. It is unchanged by the empty block. Unlike the arithmetic mean, the mean crowding is not affected by empty samples (Lloyd, 1967). Thus, contrary to the arithmetic mean, its use avoids the difficult step of delineating a domain and avoids any decisions as to what may be internal or external zeroes.

A POSSIBLE ALTERNATIVE FOR DENSITIES

The mean, the variance and Lloyd’s index of patchiness are affected by the zeroes (domain dependent). The total abundance and the mean crowding are not affected by zeroes (domain free). Basically, this corresponds to the difference between simple averages and weighted averages where zero densities get a zero weight. The following two statistics will use this simple idea when applied to density values.

Where densities are concerned, the generalization of the mean crowding is the mean fish density per
individual (Bez et al., 1997), i.e. the mean fish density around any single individual of the population. For a regular sampling grid, this is estimated by:

$$\Sigma_i(q_i \times z_i)/\Sigma_i(q_i) = \Sigma_i(Sz_i \times z_i)/\Sigma_i(Sz_i) = \Sigma z_i^2/\Sigma z_i$$

(2)

Following the same procedure as for mean crowding, we have, in our simple example (Table 1):

- 80 (i.e. 8 × 10) individuals in an area where the fish density is 8 ind. n.mi. –2;
- 20 (i.e. 2 × 10) individuals in an area where the fish density is 2 ind. n.mi. –2;
- 0 (i.e. 0 × 10) individuals in an area where the fish density is 0 ind. n.mi. –2.

The mean fish density per individual is then $(80 \times 8 + 20 \times 2 + 0) / (80 + 20) = 6.8$ ind. n.mi. –2 in both cases, that is, with or without taking into account the empty block. It is then a domain-free statistic expressed in density units.

As for the abundance, the mean density per individual is multiplied by 10 if all the densities are multiplied by 10. To compare different situations with different levels of density, it is helpful to have an index that does not change if densities are multiplied by a constant. Such an index can be defined by simply dividing the mean density per individual by the total abundance. In practice, for regular sampling schemes, this index is estimated by:

$$I_a = \Sigma z_i^2/[S \times (\Sigma z_i)^2].$$

(3)

This has the dimension of the inverse of area (for instance, n.mi. –2) and measures the probability that two fish taken at random in the population are located in the same sample (i.e. aggregated in the same tow if the sampling gear is a trawl). $I_a$ is then referred to as an index of aggregation, but should not be confused with the term 'aggregation' in spatial statistics to represent departure from Poissonness of counts such that the variance is larger than the mean. In addition, the statistics necessarily refer to the sampled population, i.e. the population within the sampled domain. The presence of 'outer' zeroes (either observed or assumed) tells us whether the extent of the entire population has been sampled and allows for a meaningful use of $I_a$.

### Spatial layout

None of the statistics presented above (i.e. the Lloyd's index of patchiness, the mean crowding, the mean density per individual and the index of aggregation) is a spatial statistic. These statistics would have the same value if the block values were distributed differently in space.

### Block size

Let us imagine blocks with a size half that of the actual size (5 n.mi.² instead of 10 n.mi.²) and the new counts of individuals distributed as in Table 2. Because the

Let us imagine two cases where the fish density is constant over space. In one case, the population extends in an area of 1 unit surface, while in the second case, the area is double.

In the first case, the index of aggregation is 1, while it is 0.5 in the second case. Because the extent of the population is smaller in the first case, the amount of individuals present in one location, relative to the total, is larger. This is to say, that taking two individuals at random in the population, they have a greater chance (in fact, double the chance) of coming from the same location in case 1 than in case 2. The population is more aggregated in the first case. This remains so if the constant density was different (larger or smaller) in, say, the second case, as in the following example:

To take into account the difference of abundance, one has to use the mean density per individual.

### ADDITIONAL REMARKS

**Zeroes are not ignored**

Whether or not zero data are used in the computation of statistics such as abundance, mean density per individual or index of aggregation does not change the results. However, the zeroes are not ignored and represent crucial information. Clearly, if a zero reading was not zero, this would change the statistic. In addition, the statistics necessarily refer to the sampled population, i.e. the population within the sampled domain. The presence of 'outer' zeroes (either observed or assumed) tells us whether the extent of the entire population has been sampled and allows for a meaningful use of $I_a$.

sampled population is unchanged, the total number of individuals (Q) is of course unchanged. This is also true for the mean fish density equal to 5 ind. n.mi.\(^{-2}\), whatever the block size. But the other statistics are affected by the block size. In particular, Lloyd's index and the index of aggregation are larger when the block size is subdivided in smaller blocks (Table 2). This can be established theoretically following Matheron (1971).

Table 2. Impact of the block size on the different statistics.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Large block</th>
<th>Comparison</th>
<th>Small block</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block size</td>
<td>10</td>
<td>&gt;</td>
<td>5</td>
</tr>
<tr>
<td>Fish counts</td>
<td>80–20</td>
<td>20–60–10–10</td>
<td></td>
</tr>
<tr>
<td>Abundance</td>
<td>100</td>
<td>=</td>
<td>100</td>
</tr>
<tr>
<td>Mean count</td>
<td>50</td>
<td>=</td>
<td>50</td>
</tr>
<tr>
<td>Variance of the counts</td>
<td>900</td>
<td>&lt;</td>
<td>4100</td>
</tr>
<tr>
<td>Lloyd's index of patchiness</td>
<td>1.34</td>
<td>&lt;</td>
<td>1.64</td>
</tr>
<tr>
<td>Mean crowding</td>
<td>67</td>
<td>&gt;</td>
<td>41</td>
</tr>
<tr>
<td>Fish densities</td>
<td>8–2</td>
<td></td>
<td>4–12–2–2</td>
</tr>
<tr>
<td>Mean density</td>
<td>5</td>
<td>=</td>
<td>5</td>
</tr>
<tr>
<td>Variance of the densities</td>
<td>9</td>
<td>&lt;</td>
<td>17</td>
</tr>
<tr>
<td>Index of aggregation</td>
<td>0.068</td>
<td>&lt;</td>
<td>0.084</td>
</tr>
<tr>
<td>Mean density per individual</td>
<td>6.8</td>
<td>&lt;</td>
<td>8.4</td>
</tr>
</tbody>
</table>

CONCLUSIONS

Lloyd's index of patchiness is domain dependent and must be used for fish counts, rather than for fish densities. Domain-free alternatives exist, such as the mean crowding (Lloyd, 1967), when applied to fish counts, or the mean density per individual or the index of aggregation (Bez and Rivoirard, 2000) when applied to fish densities.

Owing to the wording, there is a temptation to consider these tools as spatial statistics. In fact, this is not the case. The basis for developing domain-free spatial statistics has been presented in this short communication, and can be summarized by the following simple statement: arithmetic averages are affected by zeroes, whereas weighted averages where zero densities get a zero weighting are not affected by the zeroes.

REFERENCES


