SUBMITTED PAPERS

A MODIFICATION TO LEFKOVITCH'S INDEX OF SPATIAL DISTRIBUTION

K. W. DUNCAN

Department of Zoology, University of Canterbury, Christchurch

SUMMARY: A new index for contagion based on Taylor's Power Law is proposed. It is defined by $D = 4/\pi$ arctan [log (variance)/log (mean)]—1. A method of calculating confidence limits for this index and the cœfficient of dispersion is given.

INTRODUCTION

The distribution of a population of organisms throughout their habitat can be "repulsed", random or "clumped".

"Repulsed" (regular) distributions are rare in ecological situations. They result when individual organisms repulse each other so that no two organisms are in close proximity. The distribution which results is a rather regular mosaic.

Random distributions can be described by the Poisson distribution which is generated when the probability of occupance ("success" in the mathematical literature) for any one trial is independent. In other words, the organisms neither attract nor repel each other.

Clumped distributions result either when the individuals attract each other (true biological contagion) or when the distribution of the population results from the action of various factors each of which is randomly distributed. In the latter case, compound distributions such as Neyman's type-A are the most appropriate models. They have been used, for example, to fit the observed distribution of insects on leaves (Neyman 1939). For Neyman's type-A distributions to apply, the distribution of insects within or on a leaf is assumed to be random (Poisson) and the distribution of leaves is also assumed to be random (Poisson) so that the resultant overall distribution is a compound Poisson-Poisson distribution which is contagious (i.e. shows clumping). This type of "contagion" is not the same as the true biological contagion defined earlier.

Many methods of detecting contagion have been described (Bliss 1958, Bliss and Owen 1958, Greig-Smith 1964, Lefkovitch 1966, Lloyd 1967, Iwao 1969, Taylor 1961, Ibid. 1965). Almost all depend on calculating some type of index or measure from the observed data. Many of these measures depend on the variance to mean ratio (the coefficient of dispersion) which, if less than 1, indicates repulsion, if around 1 indicates randomness and if greater than 1 indicates (mathematical) contagion. It should be stressed that the existence of mathematical contagion (as indicated by a value of the variance/mean ratio greater than 1) does not necessarily show that true biological contagion is the cause of clumping because the distribution may be the result of two or more randomly distributed factors compounding to generate an overall, contagious distribution.

Unfortunately, the variance/mean ratio is not stable; it depends on the magnitude of the mean density. In situations where organisms are exhibiting an equal intensity of contagion, it has been found that the variance is related to the mean by a power function (Fracker and Brischle 1944, Taylor 1961, Ibid. 1965, Hayman and Lowe 1961). Only if the variance is linearly related to the mean would the cœfficient of dispersion be stable and independent of the magnitude of mean density.

Therefore, although the cœfficient of dispersion can be used to detect contagion, it cannot be used as a measure of the intensity of contagion.

Many attempts to overcome this difficulty have been made. One of the most promising was developed by Lefkovitch (1966) who devised an index of spatial distribution given by

$$\triangle = 1/45 \ \arctan(s^2/\sigma^2) - 1 \text{ in degrees} \quad (1)$$

or $$\triangle = 4/\pi \ \arctan(s^2/\sigma^2) - 1 \text{ in radians} \quad (2)$$

where s^2 is the sample variance and σ^2 is the parametric variance. In use, the term σ^2 is replaced by an efficient estimate for the parametric variance according to the distribution model to be tested. For some of the more common models, efficient estimates for σ^2 are:

Model:	*Estimate of* σ^2:
Poisson	m
Taylor's power law	am^b
Negative binomial	$m + m^2/k$

where m is the sample mean, b and a are cœfficients found from the regression of log s^2 on log m, and k is the exponent of the negative binomial distribution.

For the purposes of detecting contagion, the Poisson distribution is usually used as the theoretical model, since the estimates of σ^2 based on other models require considerable calculation. Furthermore, the use of this distribution gives the index the desirable property that numerical values of around zero indicate that the distribution is Poisson (random) while positive and negative values reflect clumping and repulsion respectively.

$\triangle$—index, when used with a Poisson expectation (m as an estimate of σ^2) represents a scaling of the cœfficient of dispersion so that the numerical range is $-1 - 0 - +1$. It does not stabilize the variance/mean ratio. Its values, therefore, are a function of mean density. This is the same weakness as shown by the cœfficient of dispersion (variance/mean ratio) and its derivatives.

D-INDEX

To overcome this difficulty a modified index is proposed which is defined by:

$$D = 4/\pi \ \arctan[\log s^2/\log m] - 1 \text{ (in radians)}.$$

The rationale behind this index is that the relationship between the sample variances and the sample means for a series of samples drawn from a contagious population is not linear, rather it follows a power law (Taylor 1961, Ibid. 1965) so that the ratio of log variance to log mean is constant. D-index represents an attempt to stabilize the variance/mean ratio by using the logarithms of the variance and mean. This ratio can be regarded as a tangent. Therefore, to make it more meaningful, the arctan is taken either by using tan tables 'backwards' or using published tables of arctan (Anon. 1953). The resulting value (which is in radians) is then manipulated mathmatically so that it falls within the range $-1 - +1$.

The index is not dependent on the magnitude of mean density because it is based on Taylor's power law. Thus it can be used to compare samples of unequal size.

Like $\triangle$, D takes the values

Perfect negative contagion (regular)	-1
Random (Poisson)	0
Perfect positive contagion (aggregated)	$+1$

To test the performance of the index the data collected by Howe (1950)—and used by Lefkovitch—has been reprocessed. The values for contagion as measured by $\triangle$ and D are given in Table 1.

TABLE 1. $\triangle$ *and* D *Values for Howe's (1950) Data.*

Month	Mean number of spider beetles per sack	Variance	$\triangle$*	D
1	0.6803	1.5221	0.410	0.218
2	2.3880	9.0218	0.681	0.520
3	2.6557	18.5774	0.818	0.583
4	4.9011	26.4835	0.763	0.421
5	9.5628	93.2004	0.869	0.411
6	11.2377	96.5161	0.851	0.381

* Based on a Poisson expectation.

On the basis of the △ values, Lefkovitch concluded that the population became more aggregated between the first and third month and then remained about the same. The D values, however, suggest that the aggregation reached a peak in the third month then steadily declined in spite of an increase in density.

In Lefkovitch's table (p.92) the values of △ based on a binomial, or power law, distribution do not agree with the values given by △ (based on a Poisson distribution) or D. This lack of agreement is probably due to the fact that the data for spider beetles (Table 1) is too heterogeneous to allow the fitting of a common k or power law function.

As a further test of the D-index, unequal sized samples of terrestrial amphipods were taken and analysed for aggregation using both △ and D (Table 2). The samples were taken at random with corers of 50cm² from an area of waste grassland. Further details of the study are given in an earlier paper (Duncan 1969). The interval between sampling was about six weeks.

The patern of contagion shown by the two indices is much the same for *Orchestia hurleyi* but the influence of mean density can be seen when comparing the values given by the two indices for the samples with the lowest means (samples 2 and 4). For *Orchestia patersoni*, however, the patterns given by the two indices are completely different. This possibly reflects the dependence of △ on the magnitude of the mean.

Confidence Limits for D

If α is the required confidence level, the $1-\alpha$ confidence interval for the parametric variance is given by:

$$\sigma^2 \chi^2_{(1-\alpha/2,f)} \leq fs^2 \leq \sigma^2 \chi^2_{(\alpha/2,f)}$$

where n is the sample size and $f = n-1$.
Thus

$$\frac{fs^2}{\chi^2_{(\alpha/2,f)}} \leq \sigma^2 \leq \frac{fs^2}{\chi^2_{(1-\alpha/2,f)}} \qquad (1)$$

where n is the sample size and χ^2 is read off the usual chi-square tables.

The confidence interval for the parametric mean (μ) is given by

$$m - \frac{ts}{\sqrt{n}} \leq \mu \leq m + \frac{ts}{\sqrt{n}}$$

TABLE 2. *Aggregation as Measured by* D *and* △ *in Unequal-sized Samples.*

Sampling period	Sample size	Mean number of animals/ sample (core)	Variance	Contagion as measured by: △	D
(a) *Orchestia hurleyi*					
1	9	8.778	85.728	0.8701	0.4219*
2	12	6.250	48.021	0.8351	0.4371*
3	16	8.625	50.359	0.7840	0.3600*
4	18	6.389	25.126	0.6829	0.3608*
5	20	10.950	111.147	0.8749	0.4014*
6	22	8.773	53.994	0.7949	0.3652*
7	23	13.130	130.461	0.8723	0.3809*
(b) *Orchestia patersoni*					
1	9	2.222	10.617	0.7373	0.5850
2	12	2.917	10.910	0.6674	0.4638*
3	16	3.000	12.500	0.7001	0.4776*
4	18	1.722	5.423	0.6085	0.6039*
5	20	1.850	7.228	0.6809	0.6161*
6	22	1.500	5.341	0.6514	0.6977*
7	23	4.542	34.498	0.8334	0.4858

* Significantly different from zero at a probability level of 0.05.

where t is Students—t for f degrees of freedom and s is the sample standard deviation. The confidence limits to D can be calculated using these relations.

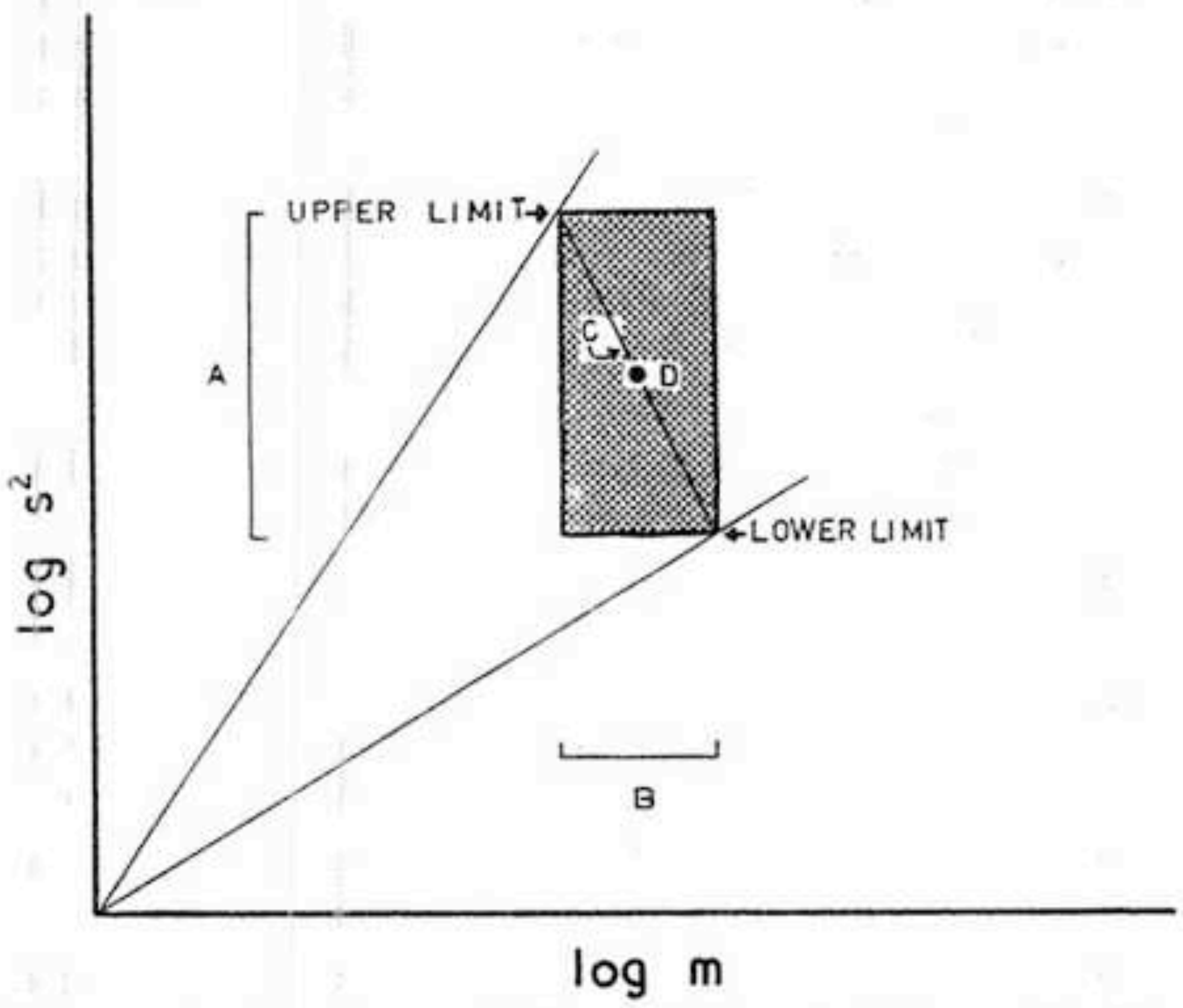

FIGURE 1. *A representation of the confidence limits to* D. *The shaded rectangle shows the confidence area.*

A—the log of the confidence interval for the variance;

B—the log of the confidence interval for the mean;

C—the line whose limits are the confidence limits for D;

D—the measured value of the index.

As shown in Figure 1, the logs of the two confidence intervals define a rectangular confidence area. The coordinates of the corners of this confidence area are:

$$[\log(m-ts/\sqrt{n}), \log(fs^2/\chi^2_{(1-\alpha/2,f)})]$$

$$[\log(m+ts/\sqrt{n}), \log(fs^2/\chi^2_{(1-\alpha/2,f)})]$$

$$[\log(m-ts/\sqrt{n}), \log(fs^2/\chi^2_{(\alpha/2,f)})]$$

$$[\log(m+ts/\sqrt{n}), \log(fs^2/\chi^2_{(\alpha/2,f)})]$$

The two extreme points which limit D are the first and last of these four. Putting these two limiting points in turn into the expression for D gives the required confidence limits.
That is:

Another approach to the problem of setting confidence limits to indices of dispersion has been made by Lefkovitch (and others). He stated that the expectation from a Poisson series is that

$$\chi^2 = \frac{fs^2}{m}$$

This is not strictly true. The expectation is that

$$\chi^2 = \frac{fs^2}{\sigma^2}$$

Since, for the Poisson distribution, m is only an estimate of σ^2, error in the mean could lead to incorrect results. Furthermore, s^2 is not an efficient estimate of σ^2 when the data is kurtotic. For the Poisson distribution, however, the ratio s^2/m can be regarded as an F-ratio with n—1 numerator and n—1 denominator degrees of freedom since both s^2 and m are independent estimates of σ^2.

Confidence intervals for the cœfficient of dispersion (CD) can be given analogous to those for D:

$$CD_{(1-\alpha/2)} = (fs^2/\chi^2_{(1-\alpha/2,f)})/(m-ts/\sqrt{n})/n$$

and $$CD_{(\alpha/2)} = (fs^2/\chi^2_{(\alpha/2,f)})/(m+ts/\sqrt{n})/n$$

Note that D-index does not work when the mean or the variance is less than 1. For such cases a modified index (D′) can be calculated using means and variances transformed by adding 1 but the values given by D′ cannot be compared directly with the values given by D. A better approach, however, is to select a sampler size which will give a distribution with m and s^2 greater than 1.

DISCUSSION

Many measures of contagion have been proposed but most suffer from one or more disadvantages. Some are dependent on the mean density while others are based on theoretical distributions whose applicability to practical situations may be

Upper limit = $4/\pi \arctan[\log(fs^2/\chi^2_{(1-\alpha/2,f)})/\log(m-ts/\sqrt{n})]-1$

Lower limit = $4/\pi \arctan[\log(fs^2/\chi^2_{(\alpha/2,f)})/\log(m+ts/\sqrt{n})]-1$

for a confidence interval of $1-\alpha$.

in doubt. Of the various indices, the exponent in Taylor's power law is the most widely applicable since it has been found to hold in a wide variety of situations. In order to calculate the exponent, however, it is necessary to take groups of samples so that the regression line of log variance on log mean can be fitted. The exponent, therefore, cannot measure contagion in a single sample. D-index has the same empirical basis as Taylor's exponent and so the index can be regarded as extending the method devised by Taylor to single samples.

REFERENCES

ANON. 1953. Table of arctan. *National Bureau of Standards Applied Maths. Series 26.* U.S. Dept. of Commerce.

DUNCAN, K. W. 1969. The ecology of two species of terrestrial Amphipoda [Crustacea: Family Talitridæ] living in waste grassland. *Pedobiologia* 9: 323-341.

FRACKER, S. B. and BRISCHLE, H. A. 1944. Measuring the local distribution of *Ribes*. *Ecology* 25: 283-303.

HAYMAN, B. I. and LOWE, A. D. 1961. The transformation of counts of the cabbage aphid (*Brevicoryne brassicae* (L.)). *New Zealand Journal of Science* 4: 271-263.

HOWE, R. W. 1950. Studies on beetles of the family Ptinidae. III A two year study of the distribution and abundance of *Ptinus tectus* Boield in a warehouse. *Bulletin of Entomological Research* 41: 371-394.

LEFKOVITCH, L. P. 1966. An index of spatial distribution. *Research in Population Ecology* 8: 89-92.

LLOYD, Monte 1967. 'Mean Crowding'. *Journal of Animal Ecology* 36: 1-30.

NEYMAN, J. 1939. On a new class of 'contagious' distributions, applicable in entomology and bacteriology. *Annals of Mathematical Statistics* 10: 35-55.

TAYLOR, L. R. 1961. Aggregation, variance and the mean. *Nature* 189: 732-735.

TAYLOR, L. R. 1965. A natural law for the spatial disposition of insects. *Proceedings of the Twelfth International Congress of Entomology* 396-397.