Tests of Randomness Based on Distance Methods
Author(s): P. Holgate
Source: *Biometrika*, Vol. 52, No. 3/4 (Dec., 1965), pp. 345-353
Published by: Oxford University Press on behalf of Biometrika Trust
Stable URL: http://www.jstor.org/stable/2333688
Accessed: 03-11-2015 22:56 UTC

Biometrika (1965), **52**, 3 *and* 4, *p.* 345
With 4 *text-figures*
Printed in Great Britain

Tests of randomness based on distance methods

BY P. HOLGATE
The Nature Conservancy, London

1. INTRODUCTION

The most familiar method of testing the hypothesis that an observed spatial distribution of points in the Euclidean plane is a realization of a Poisson point process, or in practical terminology that the points are distributed 'completely at random' is based on the Index of Dispersion. In recent years, however, statistical ecologists have shown interest in what are known as distance, or nearest neighbour methods for studying the pattern of plant distributions. These involve choosing a set of n sampling points in a fixed sampling area B according to some rule, and measuring the distances X_{ij} from the ith sample point to the jth nearest plant to it, measurements to plants outside B being permitted if necessary. The observations X_{ij} are then used in estimators of the density of the spatial point process of which the plant distribution is a realization, or in testing hypotheses about its form. Distance methods have been discussed by Skellam (1952), Strand (1953), Moore (1954), Hopkins (1954), Thompson (1956), Matérn (1959), Pielou (1959), Mountford (1961), Kendall & Moran (1963, p. 38), Keuls, Over & De Wit (1963), Persson (1964) and Holgate (1964, 1965*a*, *b*). Tests of randomness based wholly or partly on distance methods have been proposed by Skellam, Moore, Hopkins and Pielou (see also Mountford's paper). These tests are discussed in the present paper, with particular emphasis on their power to detect certain kinds of departure from randomness. In Holgate (1965*a*) two more tests are proposed and discussed.

2. PROBABILITY MODELS

Suppose first that a single sample point is chosen, and consider the random distance X_{11} $(=X)$ to the nearest plant, where the spatial distribution of plants is supposed to be a realization of a Poisson process with density λ. It is a function on the space of all possible realizations of the process, and if the sample point is chosen independently of the realization, X has the probability density function (P.D.F.) given by Skellam (1952),

$$f(x) = 2\pi\lambda x \exp(-\pi\lambda x^2). \tag{1}$$

It is convenient to transform to a new variable $Y = \pi X^2$, which consequently has P.D.F.

$$g(y) = \lambda \exp(-\lambda y). \tag{2}$$

When more than one sample point is involved there are several different possible models. For instance, if the sample points are held fixed and randomization over the space of all realizations of the Poisson process is considered, then each X_{i1} will be distributed with P.D.F. (1), but they will not be independent of each other. If, on the other hand, a fixed realization of the process is taken, i.e. attention fixed on a definite arrangement of plants, and the n sample points chosen independently, say with uniform probability density in B, then the X_{i1} will be independent but will not have P.D.F (1). Each of these models is appropriate in certain situations. The assumption implicit in much work on distance methods, that the X_{i1}

are distributed mutually independently, each with P.D.F (1) holds only if randomization over all realizations of the point process and all realizations of the choice of sampling points independently of each other is taken into account. However, the common sense idea that if the sample points are not too dense, distances from them to nearby plants will be almost independent can be formalized as a limiting result. Let $B_1 \subset B_2 \subset B_3 \subset \ldots$ be an infinite sequence of convex sampling areas whose union is the whole plane. For each realization of the Poisson process let a sequence of random variables $X_{i1}^{(1)}, X_{i1}^{(2)}, \ldots$ be defined by taking $X_{i1}^{(s)}$ to be the distance from a random point distributed with uniform probability over B_s, to the nearest plant of the (fixed) realization. Then for all realizations except for those of a set having probability zero, $X_{i1}^{(s)}$ tends in distribution to a random variable with P.D.F. (1), as $s \to \infty$. The details of the proof are omitted here. As well as distances from random points to the nearest plant, some distance methods involve measurements from randomly selected sample plants to their nearest neighbours. Although there are practical and conceptual difficulties involved, it will be assumed, following Hopkins (1954), that it is possible to choose a sample of k plants from those occurring in a given study area and for a given realization of the point process, such that each of the plants present has an equal chance of being included in the samples. Again, if an increasing sequence of sampling areas is considered as before, and the size k of the plant sample is held fixed, the distances ξ_{i1} from the ith sample plant to its nearest neighbour tend to be distributed independently, each with P.D.F (1). In the remainder of the paper, it will be assumed that the sampling area is so large that the above limiting results can be assumed to hold for practical purposes.

3. TESTS OF RANDOMNESS

To simplify notation, only distances to nearest plants will be considered from now on, and X_{i1} denoted by X_i. The following three consistent estimators of λ are available:

$$\hat{\lambda}_1 = \Sigma z_i/(mD), \quad \hat{\lambda}_2 = n/(\Sigma y_j), \quad \hat{\lambda}_3 = k/(\Sigma \eta_r).$$

z_i is the number of plants in the ith of a set of m quadrats each of area D, x_j is the distance from the jth of a set of n random sample points to the nearest plant to it, and $y_j = \pi x_j^2$, and ξ_r is the distance from the rth of a set of k random plants to its nearest neighbour, and $\eta_r = \pi \xi_r^2$. $\hat{\lambda}_2$ and $\hat{\lambda}_3$ are discussed, e.g. by Moore (1954). The tests of randomness considered here are based on ratios of pairs of estimators chosen to make them independent of the density, λ. Hopkins (1954) proposed the 'coefficient of aggregation',

$$A = \hat{\lambda}_3/\hat{\lambda}_2 = k\Sigma y_j/(n\Sigma \eta_r),$$

which is distributed as a variance ratio, $F(2n, 2k)$, on the hypothesis of randomness. (Actually, Hopkins limited his discussion to the case $k = n$, and Skellam (1954) showed that $\Sigma y_j/(\Sigma y_j + \Sigma \eta_r)$ was a beta variable, which is equivalent to the result stated above.) Pielou (1959) proposed the 'index of non-randomness'

$$\alpha = \hat{\lambda}_2/\hat{\lambda}_1 = \frac{\Sigma z_i}{mD} \Big/ \left(\frac{n}{\Sigma y_j}\right).$$

In deriving the mean and variance of α, she did not allow for variation in $\hat{\lambda}_1$, and the correct values were given by Mountford (1961),

$$\mathscr{E}\alpha = 1, \quad \operatorname{var}\alpha = \frac{1}{n} + \frac{1}{mD\lambda} + \frac{1}{mnD\lambda}.$$

Unfortunately, the variance involves the nuisance parameter λ. The likelihood of the sample of m quadrat counts and n distance measurements is

$$L = \lambda^{n+\Sigma z_i} \exp\{-\lambda(mD+\Sigma y_j)\} D^{\Sigma z_i}/(\Pi z_i!). \tag{3}$$

This shows that there is no single sufficient statistic for λ but that $(\Sigma z_i, \Sigma y_j)$ is a minimal sufficient pair. In practice the Pielou–Mountford test would be used with an estimated value of var α based either on the maximum-likelihood estimator of λ, which from (3) is

$$\hat{\lambda} = (\Sigma z_i + n)/(mD + \Sigma y_j),$$

or on the quadrat estimator $\hat{\lambda}_1$. (The choice between $\hat{\lambda}$ and $\hat{\lambda}_1$ will affect the performance of the test.) As $mD \to \infty$, var $\alpha \to n^{-1}$ and it is easy to find asymptotically similar critical regions for α. However, the assumption that mD can be taken arbitrarily large implies the assumption that λ can be determined with arbitrarily high accuracy. In this case the normal approximation can be avoided and the test introduced independently by Skellam (1952) and Moore (1954), based on the fact that $2\lambda\Sigma y_j$ is distributed as χ^2 with $2n$ D.F., used instead of α.

4(*a*). Hopkins's test against the Thomas process

The first class of alternatives to the Poisson processes to be considered is a class of point processes giving rise to the distribution introduced by Thomas (1949), which will be referred to as Thomas processes. The plants are assumed to occur in clusters, which are in turn taken to arise from a Poisson with density λ'. Each cluster contains $1+U$ plants, where U is a Poisson variable with parameter μ, and the clusters are mutually independent. The plant density is therefore $\lambda'(1+\mu)$. It will be assumed that all the plants of a given cluster are situated at the same geometrical point. Although this is a strong assumption, Skellam (1958) has pointed out that it is generally necessary for the elementary applications of the theory of contagious distributions in ecology. The practical validity of this theory depends on the results remaining approximately true when the assumptions are slightly relaxed. The 'clusters' for which $U = 0$ will consist of single plants, and will be referred to as isolated plants. As the sampling area increases to cover the whole plane, the proportion of plants—not the proportion of clusters—that are isolated will tend, except for a set of realizations of probability zero, to the value

$$\begin{aligned} p &= e^{-\mu} \Big/ \left\{ e^{-\mu} \left(1 + \frac{2\mu}{1!} + \frac{3\mu^2}{2!} + \ldots \right) \right\} \\ &= e^{-\mu}/(1+\mu). \end{aligned} \tag{4}$$

Consider a very large sampling area, and suppose that a sample of k plants is chosen so that each plant has an equal chance of being included. The probability that v of the k plants are isolated is given by the binomial probability

$$\binom{k}{v} p^v (1-p)^{k-v}, \tag{5}$$

where p is given by (4). For an isolated plant, the distance to its nearest neighbour has P.D.F. (1), while for a non-isolated plant it is zero. In studying the power of Hopkins' test against the Thomas process it is convenient to consider an equivalent form

$$A' = k/(k+nA) = \Sigma\eta_r/(\Sigma\eta_r + \Sigma y_j),$$

which on the random hypothesis is a beta variable with $(2k, 2n)$ degrees of freedom. If,

If, however, the plants arise from a Thomas process, it can be seen that conditional on v, A' is a beta variable with $(2v, 2n)$ degrees of freedom. Consequently, if u is an upper significance point for A', the probability of exceeding it on the alternative hypothesis is

$$1 - \sum_{v=0}^{k} \binom{k}{v} p^v (1-p)^{k-v} I_u(v, n), \tag{6}$$

where $I_u(v, n)$ is the incomplete beta function and the notation is that of Pearson (1934). Similarly, the probability that the test statistic will fall below the lower significance point l is

$$\sum_{v=0}^{k} \binom{k}{v} p^v (1-p)^{k-v} I_l(v, n) \tag{7}$$

and the power of the test is the sum of (6) and (7). In Fig. 1, the power of the 5 % equal-tail area version of Hopkins's test is shown against the Thomas processes with varying μ, for four sample sizes with $n = 10, 25$ and $k = 10, 25$.

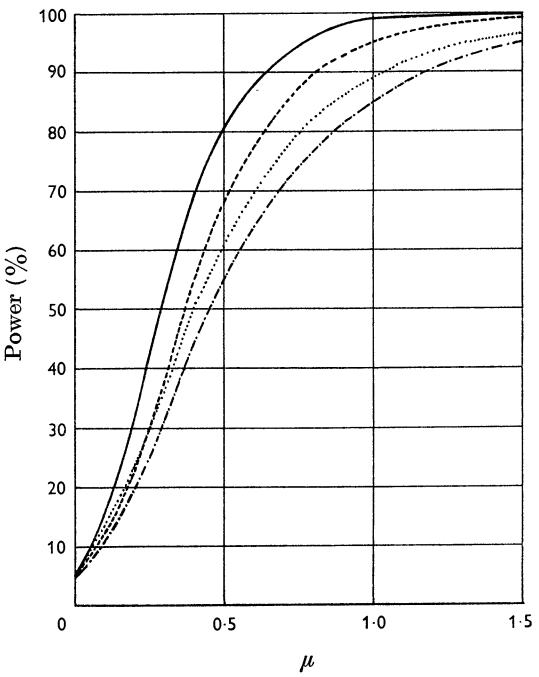

Fig. 1. The power of Hopkins's test against the Thomas process. ——, $n = k = 25$; – – – –, $n = 25$, $k = 10$; · · ·, $n = 10$, $k = 25$; –·–·–, $n = k = 10$.

4(*b*). The Skellam–Moore test against the Thomas process

This test is carried out on the assumption that the true density $\lambda'(1+\mu)$ is known. The effect of clustering is that the distance from a random point to the nearest plant is distributed as if the density were λ'. Hence for sample size n, the probability that the test statistic will exceed its upper significance point is the probability that a χ^2 variable with $2n$ degrees of freedom will exceed $(1+\mu)^{-1}$ times its upper significance point. A similar argument applies to the lower point. The sum of these probabilities is the power and, for the same version of Hopkins's test, this is plotted against μ, for sample sizes $n = 10, 25$, in Fig. 2.

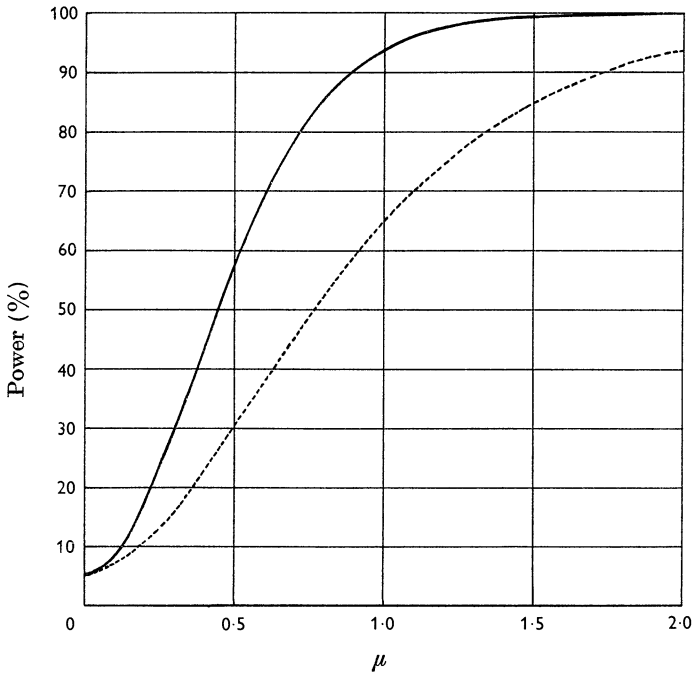

Fig. 2. The power of the Skellam–Moore test against the Thomas process. ——, $n = 25$; – – – –, $n = 10$.

5(*a*). Hopkins's test against a lattice of clusters

It is of interest to study models of vegetation communities in which different kinds of pattern occur at different spatial scales. In the class of alternatives to randomness considered in this and the following section, the plants are clustered on the smallest scale, the distribution of plants per cluster being the same as for the Thomas process. However, the clusters themselves are taken to be regularly spaced at the points of a square lattice which without loss of generality can be taken to have unit spacing. By taking $\mu = 0$, a lattice of single plants is included, but there is of course no value of μ for which the model gives a Poisson process. Consequently Figs. 3 and 4, in which the powers of the two tests against this class of alternatives are plotted, do not have the usual 'power curve' shape. The distribution of the distance from a random point to the nearest point of a square lattice has been studied by Persson (1964). The p.d.f. of $Y = \pi X^2$ in this case is

$$\left.\begin{aligned} g(y) &= 1 && (0 \leqslant y \leqslant \tfrac{1}{4}\pi), \\ &= 1 - \frac{4}{\pi}\cos^{-1}\frac{1}{2}\left(\frac{\pi}{y}\right)^{\frac{1}{2}} && (\tfrac{1}{4}\pi \leqslant y \leqslant \tfrac{1}{2}\pi). \end{aligned}\right\} \tag{8}$$

The moments of Y about zero, derived by Holgate (1965*b*), are

$$\mu_r' = \left(\frac{\pi}{4}\right)^r \sum_{s=0}^{r} \binom{r}{s} \frac{1}{(2s+1)(2r-2s+1)}.$$

In particular $\mu_1' = \frac{1}{6}\pi$, $\mu_2 = \frac{1}{90}\pi^2$. For a very large sampling area, the probability that a plant will be isolated is given by (4), and in this case the distance to its nearest neighbour is

unity. From a non-isolated plant the distance is zero. If v of the k members of the plant sample are isolated, the term $\Sigma\eta_r$ in the denominator of A has the value $v\pi$. From the moments given above, the mean and standard deviation of A, conditional on v, can be calculated and are

$$\mathscr{E}A = k/(6v), \quad \sigma(A) = k(90n)^{-\frac{1}{2}}/v. \tag{9}$$

The mean of n variables, each with P.D.F. (8) will tend quickly to normality. The normal approximation to the probability of rejecting the hypothesis of randomness, conditional on v, can be calculated using the moments (9) and then summed over v with binomial weights (5). In Fig. 3 the probability of rejecting the hypothesis of randomness using an equal-tailed 5 % test, for the alternative hypothesis of a lattice of clusters, is shown against μ for $n = 10, 25$, $k = 10, 25$.

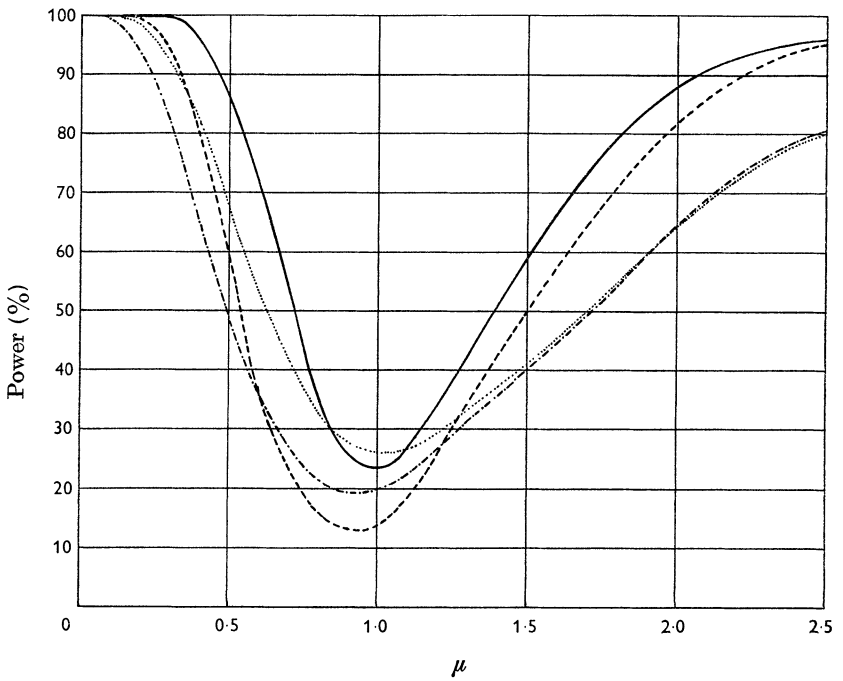

Fig. 3. The power of Hopkins's test against a square lattice of clusters. ——, $n = k = 25$; – – –, $n = 10$, $k = 25$; · · · · ·, $n = 25$, $k = 10$; –·–·–, $n = k = 10$.

5(*b*). The Skellam–Moore test against a lattice of clusters

The plant density $1+\mu$ is assumed known. The Skellam–Moore test can be re-written so that the observed mean of Y is taken as the test statistic. On the hypothesis of randomness it is distributed as $(1+\mu)^{-1}(2n)^{-1}\chi^2$, with $2n$ degrees of freedom, while on the alternative hypothesis its mean and standard deviation are

$$\mathscr{E}\overline{Y} = \tfrac{1}{6}\pi, \quad \sigma(\overline{Y}) = \pi(90n)^{-\frac{1}{2}}.$$

Again, the normal approximation to the probability of rejecting the hypothesis of randomness will be quite accurate even for moderate sample sizes; for a 5 % equal-tail test and sample sizes $n = 10, 25$, it is plotted against μ in Fig. 4.

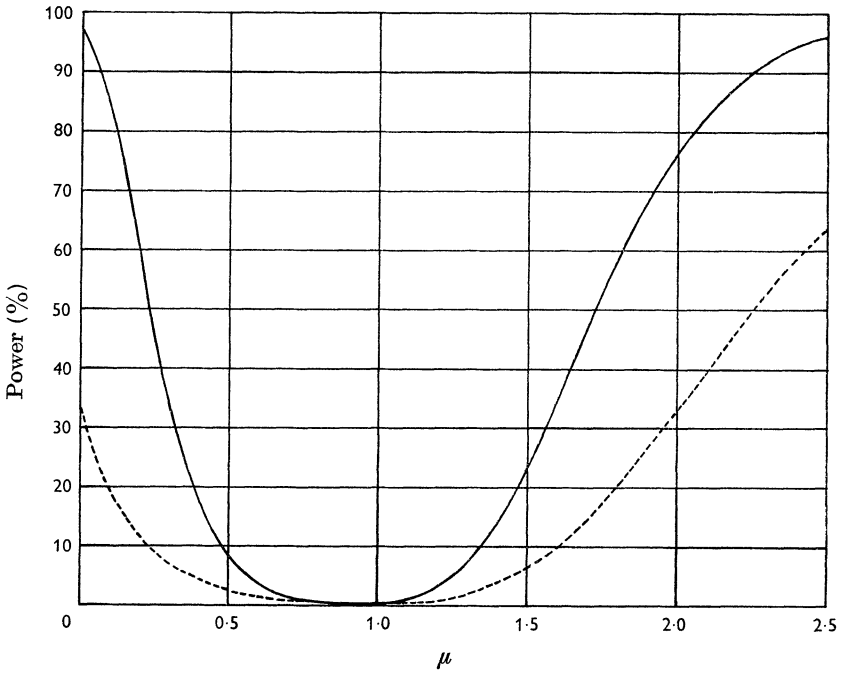

Fig. 4. The power of the Skellam–Moore test against a square lattice of clusters. ——, $n = 25$; – – –, $n = 10$.

6. Discussion

It is difficult to compare the tests of randomness studied above because they depend not only on different mathematical treatment of the data, but on different ways of collecting them. An attempt to compare estimators based on these different kinds of data has been made by Holgate (1964). Furthermore, as in all problems of randomness, the choice of alternatives against which to study the behaviour of tests is more arbitrary than in tests for parameter values. The Thomas process typifies a line of development which has been of great importance in statistical ecology and in the present context it is more convenient to deal with than a similar process leading to the Neyman Type A distribution, because it is impossible for a 'cluster' to contain no plants at all. Against these processes Hopkins's test appears to be generally more powerful than the Skellam–Moore test, despite the fact that the latter presupposes knowledge of the plant density. This can be explained by the fact that clustering simultaneously reduces average plant to plant distances and increases average point to plant distances, while the density remains unaltered.

The interest of the second alternative model has been mentioned at the beginning of § 5*a*. For values of μ close to zero, the density is close to unity, and the effect of the regular spacing is that $\mathscr{E}Y$ is much greater for a Poisson process with equal density than for the lattice. For very large μ, corresponding to very high densities, the Poisson process with the same density has an $\mathscr{E}Y$ much less than for the lattice of clusters. The probability of detecting departures from randomness is therefore high for these ranges of values of μ and the tests behave in a normal way, with increased power for larger sample sizes. On the other hand, for $\mu = 0{\cdot}9099$, the lattice of clusters has to be compared with a Poisson process with

density 1·9099 and for both these models, $\mathscr{E}Y = 0{\cdot}5236$. Moreover, a further effect of regularity of spacing, illustrated numerically in Table 1 of Holgate (1965*b*), is to reduce the variance of Y in comparison with the Poisson process. Consequently, for values of μ near 0·9099, test statistics based on Y will be very closely concentrated about their expectations, which in turn will be close to their expectations on the Poisson hypothesis. In these cases the probability of accepting the hypothesis of randomness may be considerably higher than for a truly random process. This can be seen clearly in Fig. 4. The behaviour of Hopkins's test is more complicated and has the peculiar property that, for fixed n, the power decreases as k increases. This effect is apparent from Fig. 3 and is illustrated more fully in Table 1, for

Table 1. *The power of Hopkins's test against a square lattice of clusters with* $\mu = 1{\cdot}0145$, *for varying* n, k

n \ k	5	10	25	30	60	∞
10	45·74	20·59	14·53	12·36	6·20	0·80
25	52·38	26·20	23·89	21·59	13·51	0·87

the particular value $\mu = 1{\cdot}0145$, corresponding to $p = 0{\cdot}18$. As $k \to \infty$, the random proportion of isolated plants which appears in the denominator of A converges to a constant and the behaviour of the test is determined by the numerator which is proportional to the observed mean of Y. However, an examination of the individual terms in the sum described just below expression (9), shows that for small sample sizes there is a considerable probability that A will fall in the critical region, which appears to be accounted for by (i) the randomness of the denominator, which is proportional to a binomial variable with $p = 0{\cdot}18$, and (ii) the probability of very small values of the denominator resulting in very high values of A. There is in particular a probability $(1-p)^k$ that $A = \infty$. In other words for values of μ where the lattice of clusters is confusable with a Poisson process, the probabilities of detecting it by Hopkins's test is asymptotically negligible and this describes, in a sense, the true properties of the test; however, for small k the probability is higher, due to the difference between the small sample and the asymptotic behaviour of the test.

I am grateful to Mr D. Spalding and Mr K. Lakhani for assistance in computation.

REFERENCES

HOLGATE, P. (1964). The efficiency of nearest neighbour estimators. *Biometrics*, **20**, 647–9.
HOLGATE, P. (1965*a*). Some new tests of randomness. *J. Ecol.* **53**, 261–6.
HOLGATE, P. (1965*b*). The distance from a random point to the nearest point of a closely packed lattice. *Biometrika*, **52**, 261–3.
HOPKINS, B. (1954). A new method of determining the type of distribution of plant individuals. *Ann. Bot.* **18**, 213–26.
KENDALL, M. G. & MORAN, P. A. P. (1963). *Geometrical Probability*. London: Charles Griffin and Co.
KEULS, M., OVER, H. J. & DE WIT, C. T. (1963). The distance method for estimating densities. *Stat. Neerl.* **17**, 71–91.
MATÉRN, B. (1959). Några tillämpningar av teorin for geometriska sannolikheter. *Svenska skogsvårdsföreningens tidskr.* **3**, 453–8.
MOORE, P. G. (1954). Spacing in plant populations. *Ecology*, **35**, 222–7.
MOUNTFORD, M. D. (1961). On E. C. Pielou's Index of Non-Randomness. *J. Ecol.* **49**, 271–5.
PEARSON, K. (1934). *Tables of the Incomplete Beta Function*. Cambridge University Press.
PERSSON, O. (1964). Distance methods. *Studia Forestalia Suecica*, Nr. **15**, 68 pp.

PIELOU, E. C. (1959). The use of point to plant distances in the study of the pattern of plant populations. *J. Ecol.* **47**, 607–13.

SKELLAM, J. G. (1952). Studies in statistical ecology. I. Spatial pattern. *Biometrika*, **39**, 346–62.

SKELLAM, J. G. (1954). Appendix to Hopkins (1954).

SKELLAM, J. G. (1958). On the derivation and applicability of Neyman's Type A distribution. *Biometrika*, **45**, 32–6.

STRAND, L. (1953). Mål for fordelingen av individer over et område. *Medd. fra det Norske skogsforsøksvesen* **12**, 195–207.

THOMAS, M. (1949). A generalisation of Poisson's binomial limit for use in ecology. *Biometrika*, **36**, 18–25.

THOMPSON, H. R. (1956). Distribution of distance to *n*th nearest neighbour in a population of randomly distributed individuals. *Ecology*, **37**, 391–4.